

Asymmetric Reference-dependent Reciprocity, Downward Wage Rigidity, and the Employment Contract*

Alex Dickson and Marco Fongoni[†]

Department of Economics, University of Strathclyde, Glasgow, UK

November 13, 2018

Abstract

We develop a model of asymmetric reciprocity and optimal wage setting based on contractual incompleteness, fairness, and reference dependence and loss aversion in the evaluation of wages by workers. The model establishes a positive wage-effort relationship capturing a worker's 'asymmetric reference-dependent reciprocity', in which loss aversion implies negative reciprocity is stronger than positive reciprocity. Our theory provides an explanation for the observed asymmetry and dynamics of workers' reciprocity and establishes a micro-foundation for downward wage rigidity, the implications of which shed new light on a forward-looking firm's optimal wage setting and hiring decisions.

JEL Codes: D91, E70, J30, J41.

Keywords: reference dependence, loss aversion, asymmetric reciprocity, downward wage rigidity.

*A previous version of this paper circulated under the title "A Theory of Wage Setting Behaviour".

[†]Corresponding author: e-mail: marco.fongoni@strath.ac.uk.

1 Introduction

There is an emerging consensus in the literature that behavioural concerns such as fairness, workers' morale and reciprocity influence firms' wage setting behaviour (Fehr et al., 2009). These aspects are also considered to be the key behavioural forces that underlie the observation of downward wage rigidity: compensation managers may refrain from cutting wages following adverse economic conditions if they believe this will negatively affect workers' morale and effort (Bewley, 2007). Inspired by these ideas, we develop a theory of asymmetric reciprocity and downward wage rigidity based on contractual incompleteness, fairness, and reference dependence and loss aversion in the evaluation of wage contracts by workers.

The theory sheds light on the sources of asymmetry and dynamics of workers' reciprocity documented by the empirical literature; and on its implications for optimal wage setting and the employment contract. In particular, the paper makes the following contributions: i) it offers a psychological foundation—based on loss aversion and reference wage adaptation—for the observed asymmetry, and temporary nature of, workers' reciprocity; ii) it provides a transparent, and plausible, theoretical micro-foundation for *dynamic* downward wage rigidity; iii) it sheds new light on the implications of 'asymmetric reference-dependent reciprocity' and wage rigidity for a forward-looking firm's optimal wage setting and hiring decisions. The paper contributes to a large body of literature that has considered the implications of reciprocity, loss aversion, and downward wage rigidity in labour markets¹ by developing a tractable microeconomic model that allows a rigorous analysis of the asymmetries and irreversibility of wage and effort dynamics, and of their implications for the nature of the employment contract. We believe that gaining a deeper understanding of the incentives driving workers' and firms' behaviour when engaged in employment relationships is also particularly important in light of recent advances in the macroeconomic theory of labour market fluctuations (Elsby et al., 2015).

¹This literature spans from the first efficiency wage models developed by Akerlof (1982) and Akerlof and Yellen (1990), to the model of Bhaskar (1990) and to more recent applications of fairness and reciprocity in labour markets, such as, Danthine and Kurmann (2007), Eliaz and Spiegel (2014), and Benjamin (2015). In Section 4 we provide an extensive discussion of this literature, [which, since the first version of our paper appeared in July 2015 (SIRE Discussion Paper SIRE-DP-2015-57), has been growing fast (see for instance Ahrens et al. (2015), Kurmann and McEntarfer (2017), Sliwka and Werner (2017), Kaur (2018) and Macera and te Velde (2018)).]

The basic premise of our theory is that there is contractual incompleteness over effort (Williamson, 1985); and that workers evaluate wage contracts relative to a reference ‘fair’ wage. Central to our model is the inclusion of a ‘morale function’ in the worker’s payoff, which consequently exhibits both positive and negative reciprocity stemming from their reference-dependent preferences: a wage above the reference wage increases morale and triggers supra-normal effort; a wage below the reference wage reduces morale and triggers sub-normal effort. This establishes a positive wage-effort relationship where, if the worker is loss averse, negative reciprocity is stronger than positive reciprocity. To the best of our knowledge our paper is the first that formally derives a link between reference dependence, loss aversion and the asymmetric nature of reciprocity from a worker’s optimal behaviour. Although the relationship between loss aversion and negative reciprocity has already been *conjectured* in the empirical literature (see e.g. Fehr et al. (2009)), we think that rigorously formalising this idea will not only enable us to transparently analyse the implications of loss aversion for optimal wage setting, but will also allow us to derive new testable hypotheses. In addition we show that if a worker uses the past wage as their reference for fairness, any reciprocity response triggered by an initial wage change will eventually disappear—reciprocity in our model is a temporary phenomenon. This implication is consistent with evidence on the dynamics of workers’ reciprocity and supports the interpretation according to which reciprocity is temporary due to dynamic adaptation of the reference wage (see, for instance, Gneezy and List (2006), Mas (2006), and Sliwka and Werner (2017)).

In the second main contribution of the paper we explore the implications of the asymmetry and dynamics of reciprocity just discussed in a two-period employment relationship in which the evolution of the job-match productivity is uncertain. First, we show that in a situation in which the firm is facing a negative shock and has an incentive to decrease the worker’s wage, consideration of the relatively large impact of negative reciprocity on output gives rise to downward wage rigidity for a range of negative shocks.² Importantly, this is not a static result: due to the worker’s adaptation to wage increases during periods

²Downward (nominal and real) wage rigidity is the tendency of wages to not fall during recessions. It has been widely documented in the empirical literature by looking at wage change distributions, which exhibit a high incidence of wage freezes with wage cuts less frequent than wage increases (see e.g. Nickell and Quintini (2003), Fehr and Goette (2005), Dickens et al. (2007)). Evidence that firms avoid cutting wages during recessions has also been reported by several field surveys (see, for instance, Campbell and Kamlani (1997), Bewley (1999), Agell and Benmarker (2007), Babecký et al. (2010)).

of expansion, the firm will face the trade-off between wage cuts and negative reciprocity whenever, in any subsequent period, it is confronted with negative shocks. Downward wage rigidity may arise even at wage levels substantially higher than those with which the employment relationship had initially started. As we discuss in our analysis we identify an asymmetric adjustment cost around the past wage as a necessary ingredient for models of dynamic downward wage rigidity (such as Elsby (2009), Holden and Wulfsberg (2009) and Kaur (2018)). In our model these features are captured by the worker’s dynamic adaptation of the reference wage, and the relatively large cost to the firm of negative reciprocity (that stems from loss aversion). As such our paper contributes to this literature by providing a plausible and transparent micro-foundation for dynamic downward wage rigidity.

Moreover, we study how a forward-looking firm’s expectation of negative reciprocity and downward wage rigidity can affect the determination of the optimal hiring wage. This analysis is important in understanding whether the expectation of future downward wage rigidity also implies a more compressed wage growth throughout the employment relationship. While the literature on this subject has established that such an expectation leads to compression of wage increases (see Elsby (2009) and Benigno and Ricci (2011)), in contrast we show that this prediction is not robust in our model: a firm that expects to be more constrained by downward wage rigidity in the future is also facing a greater *ex ante* probability of subsequently laying off the worker. This reduces the expected duration of the employment relationship, and will therefore partially offset the incentive to compress wage increases to keep their worker’s wage and reference wage low in the future.

Finally, we explore how these considerations affect the firm’s hiring decisions through the influence on the firm’s expected value of the employment relationship. We find that independently of whether the initial hiring wage is compressed by the firm, the expectation of downward wage rigidity—and the anticipation of stronger negative reciprocity—unambiguously reduce the value of a new employment relationship, implying that the firm will hire less on average. This result contributes to the literature concerned with the effects of incumbent workers’ downward wage rigidity on job creation (see the discussion in Elsby et al. (2016)) and suggests that expected rigidities in the wage of existing/incumbent workers can negatively affect firms’ incentives to hire.

The outline of the paper is as follows. We set out our model of asymmetric reference-dependent reciprocity and optimal wage setting in Section 2. In Section 3 we explore the properties of the model for wage and effort dynamics, and we study their implications for the nature of the employment contract. Section 4 provides a discussion of the related literature and highlights our main contributions, and Section 5 offers some concluding remarks. All proofs are contained in the appendix.

2 Basic Set-Up

We begin by considering an established worker-firm employment relationship for a single employment period to illustrate the mechanisms at work. We assume a setting of complete information. The worker is assumed to be reference dependent and loss averse: they evaluate wage contracts in relation to a reference ‘fair’ wage, which captures their perception of fairness and is taken as exogenous for the purpose of this section. A wage below the reference wage is perceived as *unfair*, while a wage above is perceived as a *gift*. At the start of the employment period the firm learns the (non-negative) match productivity and the worker’s reference wage, and subsequently decides on the profit-maximising wage. After observing the wage and evaluating it in relation to their reference wage, the worker decides on the utility-maximising level of effort that generates output for the firm. Payoffs are then realised.

Wage setting is thus formalised as a sequential-move game in which the firm (the first mover) makes a take-it-or-leave-it wage offer to the worker (the second mover). Since the worker’s belief of what should be a ‘fair’ wage is independent of the firm’s actions, and the firm is assumed to be motivated only by profit, the game can be solved by backward induction.³

³In contrast to more general applications of ‘psychological game theory’ to intentions-based reciprocity (e.g., Rabin, 1993; Dufwenberg and Kirchsteiger, 2004), in our model reciprocity is only on the side of the worker (as is also the case in Danthine and Kurmann (2007) and Benjamin (2015) in the labour market literature); and the worker’s perception of fairness is determined by a comparison of the wage to a wage they consider as fair—given by the reference wage—that is independent of the actions the firm could have taken had it not paid its chosen wage. As such, we do not need to consider beliefs, which would be necessary if we construct the fair wage from what could otherwise have been chosen. Nevertheless, the firm’s choice of the wage conveys intention from the worker’s perspective, and the firm’s ‘kindness’ or ‘unkindness’ is judged by the worker in terms of whether the wage offer is above, equal, or below the reference wage. Whilst this is a straightforward conception of reciprocity, it is rich enough to allow us to capture many salient features of the employment relationship, and allows us to use simple backward induction to solve

2.1 Payoffs

We let e denote effort of the worker, w the wage paid, r the reference wage, and q the match productivity. The instantaneous profit function of the firm is given by

$$\pi(w; q, e) = y(q, e) - s(w), \quad (1)$$

where y is the per-worker output (the price of which is normalised to one) and s is the per-worker cost of production. We make the following assumptions:

F1. $s' > 0$ and $s'' = 0$.

F2. $y_e, y_q > 0$, $y_{ee}, y_{qq} \leq 0$ and $y_{qe} > 0$.

We specify the worker's preferences by an additively separable utility function

$$u(e; w, r) = m(w) - d(e) + M(e; w, r), \quad (2)$$

where m captures the worker's evaluation of the wage; d represents the worker's intrinsic psychological *net cost* of productive activity;⁴ and M is the 'morale function' that depends on the worker's evaluation of the wage in relation to the reference wage:

$$M(e; w, r) \equiv e \cdot n(w|r). \quad (3)$$

We assume that $n(w|r) \equiv \mu(m(w) - m(r))$ where μ is a gain-loss value function that exhibits loss aversion in the spirit of Kahneman and Tversky (1979) and Tversky and Kahneman (1991).

We impose the following assumptions:

W1. $m' > 0$ and $m'' < 0$.

W2. $d'(0) < 0$ and $d'' > 0$.

W3. μ is piecewise-linear with $\mu(0) = 0$ and $\mu'(-x)/\mu'(x) \equiv \lambda \geq 1$ for any $x > 0$.

the game.

⁴For instance: $d(e) = c(e) - b(e)$, where c and b are respectively the physical/psychological cost and benefit of effort.

Assumption W2 implies that $u(e; w, r)$ is strictly concave in e , so that for each wage and reference wage combination there is a unique utility-maximising level of effort (and also that ‘normal’ effort—when the wage is equal to the reference wage—is positive, as we subsequently discuss). Assumption W3 closely resembles the assumptions of Kőszegi and Rabin (2006) over the properties of their universal gain-loss function, except we do not have diminishing sensitivity. It follows that the gain-loss utility takes the form:

$$n(w|r) = \begin{cases} \eta[m(w) - m(r)] & \text{if } w \geq r \text{ and} \\ \lambda\eta[m(w) - m(r)] & \text{if } w < r, \end{cases} \quad (4)$$

where $\eta > 0$ is a parameter that captures the importance of gain-loss utility and $\lambda \geq 1$ represents the worker’s degree of loss aversion.

The morale function in (3) captures the psychological cost/benefit of productive effort associated with the worker’s perception of fairness. If the wage exceeds the reference wage (it is perceived as a gift) the worker gains some additional benefit of productive effort and an increase in effort (a gift to the firm) will increase utility. If the wage falls short of the reference wage (it is perceived as unfair) there is a psychological cost of productive effort and a reduction in effort (an ‘unkind’ action towards the firm) increases utility. As such, the morale function implies the worker’s payoff exhibits reciprocity, and since morale is linked to loss aversion, negative reciprocity is stronger than positive reciprocity.

2.2 The worker’s choice of effort

Given a reference wage r and a wage offer w , the worker will seek to

$$\max_{e \geq 0} m(w) - d(e) + en(w|r).$$

The necessary (and under our assumptions sufficient) first-order condition is

$$-d'(e) + n(w|r) \leq 0, \quad (5)$$

in which the inequality is replaced with an equality if $e > 0$. To save on notational complexity, we henceforth assume an interior solution⁵ in which the optimal effort is given by

$$\tilde{e}(w, r, \lambda) = d'^{-1}(n(w|r)) = \begin{cases} d'^{-1}(\eta[m(w) - m(r)]) \equiv \tilde{e}(w, r)^+ & \text{if } w > r \\ d'^{-1}(0) \equiv \tilde{e}^n & \text{if } w = r \\ d'^{-1}(\lambda\eta[m(w) - m(r)]) \equiv \tilde{e}(w, r, \lambda)^- & \text{if } w < r. \end{cases} \quad (6)$$

When $w = r$ the morale function is zero and the worker's utility is maximised by the value of effort such that $d'(e) = 0$, referred to as 'normal' effort and denoted \tilde{e}^n , which is positive (due to the inclusion of a net cost of productive activity with the properties imposed in Assumption W2) and independent of the wage. This is consistent with the idea that workers perceive positive satisfaction from engaging with productive activity.⁶ If the worker is paid a wage above their reference wage, they will positively reciprocate this gift with 'supra-normal' effort $\tilde{e}(w, r)^+ > \tilde{e}^n$; while if the wage is set below their reference wage, they will negatively reciprocate this unfair wage by exerting 'sub-normal' effort $\tilde{e}(w, r, \lambda)^- < \tilde{e}^n$. The properties of the optimal effort function are summarised in the following theorem.

Theorem 1. *For a given r , $\tilde{e}(w, r, \lambda)$ is a continuous function of w with $\tilde{e}_w(w, r, \lambda) > 0$ and $\tilde{e}_{ww}(w, r, \lambda) < 0$ for all $w \neq r$. Moreover,⁷*

$$\lim_{\epsilon \rightarrow 0} \tilde{e}_w(r - \epsilon, r, \lambda)^- = \lambda \lim_{\epsilon \rightarrow 0} \tilde{e}_w(r + \epsilon, r)^+,$$

so the effort function has a kink at $w = r$ if $\lambda > 1$. For a given w , optimal effort $\tilde{e}(w, r, \lambda)$ is a continuous function of r with $\tilde{e}_r(w, r, \lambda) < 0$ for all $w \neq r$. Finally, whilst optimal effort above the reference wage is independent of λ , for all $w < r$: $\tilde{e}_\lambda(w, r, \lambda) < 0$ and $\tilde{e}_{w\lambda}(w, r, \lambda) > 0$.

⁵In the proof of Theorem 1 we give a sufficient condition for this to be the case.

⁶Inspired by the findings reported in Bewley (2007)—that it is not wage levels but changes in wages that influence effort—normal effort should be a non-pecuniary concept and is therefore modelled as being independent of the wage. See, for example, the discussion in Altmann et al. (2014, Appendix). A similar assumption is also considered by Sliwka and Werner (2017), Kaur (2018) and Macera and te Velde (2018).

⁷Throughout the paper, where sequences of ϵ are considered over which limits are taken we specify that $\{\epsilon_n\}_{n=1}^\infty \subset \mathbb{R}_+$, meaning that where the wage is specified to be $r - \epsilon$ and we take the limit as $\epsilon \rightarrow 0$ we consider the wage increasing to the reference wage, and likewise when the wage is specified to be $r + \epsilon$ and we take $\epsilon \rightarrow 0$ we consider the wage decreasing to the reference wage.

This relationship between effort and the wage is illustrated in Figure 1.

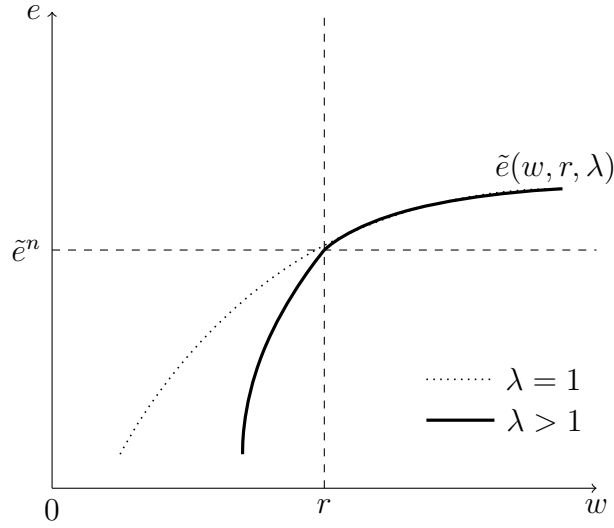


Figure 1:
Asymmetric reference-dependent reciprocity

The asymmetric nature of effort responses has the particular implication that for changes in the wage from an initial wage equal to the reference wage, the effect of negative reciprocity that results from a wage cut will be greater than the effect of positive reciprocity resulting from a wage increase. The extent of this ‘asymmetric reference-dependent reciprocity’ depends on the worker’s degree of loss aversion. Indeed, if a worker is not loss averse ($\lambda = 1$), reciprocity is symmetric.

Our derived wage-effort relationship is consistent with the large body of evidence documenting the asymmetric nature of workers’ reciprocity in response to wage changes (see, for instance, the anthropological evidence documented in Campbell and Kamlani (1997) and Bewley (1999); the field experiments of Kube et al. (2013) and Cohn et al. (2014); and the related literature surveyed in Bewley (2007), Fehr et al. (2009) and Malmendier et al. (2014)). Moreover, our model formally identifies loss aversion as the psychological foundation for why negative reciprocity is stronger, therefore providing a micro-foundation for reduced-form effort/production functions exhibiting asymmetric reciprocity that are commonly assumed in the literature (e.g. Elsby (2009), Eliaz and Spiegel (2014) and Kaur (2018)) but that are not explicitly modelled.

2.3 The firm's wage setting rule

Next we consider the firm's problem in setting the wage given that it anticipates the behaviour of the worker. After observing the worker's reference wage r and match productivity q , the firm will seek to maximise its profit given that the worker's effort is determined as in (6). As such, the firm's problem is to

$$\max_{w \geq 0} y(q, \tilde{e}(w, r, \lambda)) - s(w).$$

We showed in Theorem 1 that the worker's optimal effort function $\tilde{e}(w, r, \lambda)$ is continuous in the wage, but that there is a kink at $w = r$ for a loss averse worker. The fact that the firm's profit function is continuous and concave in w allows us to derive the optimal wage setting rule, accounting for the kink in the profit function at $w = r$.⁸ For $w \neq r$ the optimal wage is characterised by the following first-order condition

$$y_e(q, \tilde{e}(w, r, \lambda))\tilde{e}_w(w, r, \lambda) - s'(w) \leq 0, \quad (7)$$

where the inequality is replaced with an equality if $w > 0$. Again to save on notation, we henceforth assume an interior solution.⁹ The first term in (7) captures the marginal product of labour induced by a wage change, while the second term captures the related per-worker marginal cost. The resulting optimal wage setting rule is characterised by two productivity thresholds: a lower threshold q^l , which is such that if $q < q^l$ then profit is maximised where the marginal product of labour equals the marginal cost at a wage strictly below the reference wage; and an upper threshold q^u , which is such that if $q > q^u$ then profit is maximised by equating the marginal product of labour with the marginal cost at a wage exceeding the reference wage. Instead, if $q^l \leq q \leq q^u$ profit will be maximised at the kink, i.e. where $w = r$. If the match productivity falls below a reservation threshold

⁸For $w \neq r$ concavity of profit requires $y_{ee}[\tilde{e}_w]^2 + y_e\tilde{e}_{ww} < 0$. This follows from our assumptions and the conclusion of Theorem 1 that for $w \neq r$ $\tilde{e}_{ww} < 0$. To ensure concavity over the entire domain we need to ensure that at $w = r$ the marginal profit reduces, which follows again from the conclusions of Theorem 1 that imply $\lim_{\epsilon \rightarrow 0} \tilde{e}_w(r - \epsilon, r, \lambda)^- \geq \lim_{\epsilon \rightarrow 0} \tilde{e}_w(r + \epsilon, r, \lambda)^+$, where the inequality is strict if $\lambda > 1$.

⁹This requires the marginal product of labour to be sufficiently high when effort is at its lowest, so that $y_e(q, \tilde{e}(0, r, \lambda))\tilde{e}_w(0, r, \lambda) - s'(0) > 0$.

$\underline{q}(r, \lambda)$, implicitly defined by the zero-profit condition:

$$\pi(\tilde{w}(r, q, \lambda); q, \tilde{e}(\tilde{w}(r, q, \lambda), r, \lambda)) = 0$$

the employment relationship will be terminated.

These properties are summarised in the following theorem.

Theorem 2. *The optimal wage $\tilde{w}(r, q, \lambda)$ is a continuous function of q and r and is given by*

$$\tilde{w}(r, q, \lambda) = \begin{cases} \tilde{w}(r, q)^+ & \text{if } q > q^u(r) \\ r & \text{if } q^l(r, \lambda) \leq q \leq q^u(r) \\ \tilde{w}(r, q, \lambda)^- & \text{if } q < q^l(r, \lambda), \end{cases} \quad (8)$$

where $q^l(r, \lambda)$ and $q^u(r)$ are respectively characterised by the value of q such that

$$\begin{aligned} \lim_{\epsilon \rightarrow 0} y_\epsilon(q, \tilde{e}(r - \epsilon, r, \lambda))\tilde{e}_w(r - \epsilon, r, \lambda) - s'(r - \epsilon) &= 0; \text{ and} \\ \lim_{\epsilon \rightarrow 0} y_\epsilon(q, \tilde{e}(r + \epsilon, r, \lambda))\tilde{e}_w(r + \epsilon, r, \lambda) - s'(r + \epsilon) &= 0; \end{aligned}$$

(all singletons), so long as $q \geq \underline{q}(r, \lambda)$.

The optimal wage $\tilde{w}(r, q)^{+(-)}$ ($>$ ($<$) r) is implicitly defined by (7). Moreover:

- a) $\tilde{w}_q(r, q, \lambda) > 0$ for all $q \in [\underline{q}(r, \lambda), \infty) \setminus [q^l(r, \lambda), q^u(r)]$;
- b) $\tilde{w}_r(r, q, \lambda) > 0$ for all $[\underline{q}(r, \lambda), \infty)$; and
- c) $\tilde{w}_\lambda(r, q, \lambda) > 0$ for all $[\underline{q}(r, \lambda), q^l(r, \lambda))$.

In addition, $q^l(r, \lambda) < q^u(r)$ for all $\lambda > 1$ and if $\lambda = 1$, $q^l(r, 1) = q^u(r)$, and

- d) $q_\lambda^l(r, \lambda) < 0$, $q_r^l(r, \lambda) > 0$; and
- e) $q^{w'}(r) > 0$.

The reservation productivity has the properties that $\underline{q}_r(r, \lambda) > 0$ and $\underline{q}_\lambda(r, \lambda) \geq 0$, where the final inequality is strict if $\underline{q}(r, \lambda) < q^l(r, \lambda)$.

The main features of the firm's optimal wage when facing a loss averse worker are illustrated in Figure 2.

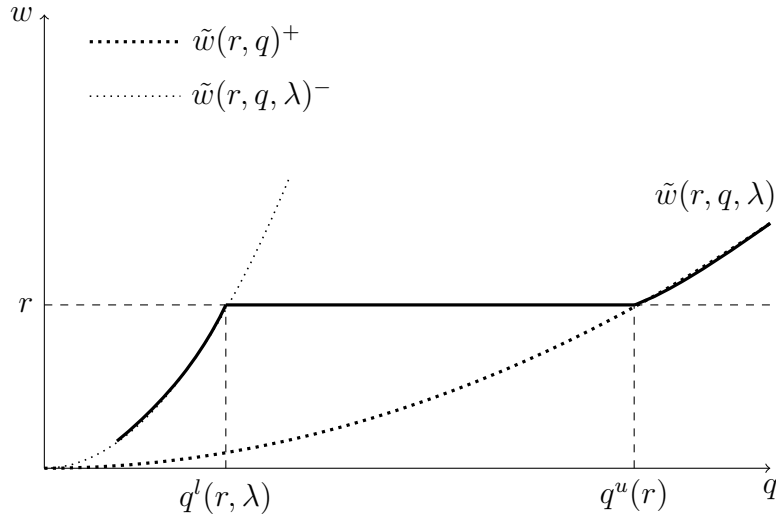


Figure 2:
The wage setting rule

The optimal wage is non-decreasing in the match productivity and there is a range of match productivity within which the wage is not adjusted. We refer to this as the ‘range of rigidity’, which is non-empty if the worker is loss averse as in this region the benefit of reducing the wage will be offset by the cost generated by the worker’s negative reciprocity response. Note that a firm employing a worker with a greater degree of loss aversion will be less willing to suffer the relatively high cost of negative reciprocity for a given match productivity (i.e. the range of rigidity becomes larger as $q_\lambda^l(r, \lambda) < 0$), and if q is low enough that the firm wishes to pay $\tilde{w}(r, q, \lambda)^- (< r)$, it will have an incentive to attenuate some of the effect from the stronger negative reciprocity, by paying a higher wage: $\tilde{w}_\lambda(r, q, \lambda)^- > 0$. Negative reciprocity not only tempers the firm’s incentive to cut the wage, it also reduces the extent to which the wage is cut.¹⁰ Finally, the more loss averse a worker is, the higher is the reservation productivity that the firm requires from the employment relationship for it to be profitable: $\underline{q}_\lambda(r, \lambda) > 0$, since a higher λ reduces per-worker output and increases the per-worker labour cost.

¹⁰This implication has been theoretically derived, and empirically corroborated, in a model of firm-level wage bargaining by Holden and Wulfsberg (2014), who show that “even if the wage is cut, the resulting wage will be higher than if the wage-setting process had been completely flexible”. In contrast with their theoretical model, our theory attributes this result to the worker’s extent of negative reciprocity.

3 Adaptation, Loss Aversion and the Employment Contract

We now turn to explore the dynamic implications of asymmetric reciprocity and reference dependence for optimal wage setting and hiring in a two-period dynamic environment. In so doing we introduce uncertainty around the evolution of the match productivity and, inspired by the literature that suggests reference points are influenced by previous contractual arrangements, we model the reference wage as being endogenously determined by the past wage (see Section 4 for a discussion of this literature). To capture these features, we impose the following two assumptions:

- D1.** The match productivity q_t follows a Markov process described by the cumulative distribution function $F(q_1|q_0)$, q_0 given.
- D2.** The worker's reference wage evolves according to the adaptation rule: $r_1 = w_0$, r_0 given.

The timing of the model is as follows. At the beginning of the first employment period, the match productivity q_0 and the exogenously-given reference wage r_0 are observed. The firm then decides whether to offer a wage contract to the worker and start the employment relationship. We assume that any offer is accepted by the worker.¹¹ If an employment relationship is established, then at the beginning of the second employment period the match productivity changes stochastically to a new value q_1 , and the worker adapts their reference wage to the wage paid in the initial period of employment. After observing q_1 , and inferring the worker's new reference wage, the firm considers whether it wants to continue the employment relationship and, if so, whether to adjust the wage in light of the change in match productivity.

The forward-looking firm therefore faces a two-period dynamic optimisation problem

¹¹For clarity of exposition, we choose not to explicitly model reservation wages; in the Appendix we show that it is straightforward to impose a condition on the exogenous variables of the model such that the worker's participation constraint is never binding which justifies this approach. That said, reservation wages might be relevant in a richer macroeconomic framework in which the worker's initial reference wage is endogenous, and, in particular, dependent on the state of the labour market.

Consistent with Assumptions F1-F2 and W1-W3 of Section 2, consider the following functional forms: per-worker output $y(q, e) = qe$; per-worker labour cost $s(w) = w$; worker's utility from the wage $m(w) = \log w$; and their net cost of productive activity $d(e) = e^2/2 - be$, with $b > 0$. Denoting $\tilde{e}(w_t, r_t, \lambda) = \tilde{e}_t$ and $\tilde{w}(r_t, q_t, \lambda) = \tilde{w}_t$, in each period $t = \{0, 1\}$ the worker's optimal effort function and the firm's optimal wage take the following simple forms:

$$\tilde{e}_t = \begin{cases} \tilde{e}^n + \eta[\log w_t - \log r_t] & \text{if } w_t > r_t \\ \tilde{e}^n & \text{if } w_t = r_t \\ \tilde{e}^n - \lambda\eta[\log r_t - \log w_t] & \text{if } w_t < r_t; \end{cases} \quad \tilde{w}_t = \begin{cases} \eta q_t & \text{if } q_t > q^u(r_t) \\ r_t & \text{if } q^l(r_t, \lambda) \leq q_t \leq q^u(r_t) \\ \lambda\eta q_t & \text{if } q_t < q^l(r_t, \lambda); \end{cases}$$

where $\tilde{e}^n = b$;

$$q^l(r_t, \lambda) = \frac{r_t}{\lambda\eta} \quad \text{and} \quad q^u(r_t) = \frac{r_t}{\eta};$$

and we assume that in both the initial and subsequent employment periods the match is profitable.

Consider a worker characterised by a relatively high match productivity $q_0 > q^u(r_0)$ such that the firm will find it profitable to hire them and pay a wage gift $\tilde{w}_0 = \eta q_0 (> r_0)$, which is positively reciprocated by supra-normal effort $\tilde{e}_0 > \tilde{e}^n$ in the first employment period. This is illustrated in Figure 4a below. As the employment relationship passes into the second employment period, the worker adjusts their feelings of entitlement, adapting their reference wage to their initial wage: $r_1 = \tilde{w}_0 = \eta q_0$. This 'shifts' the wage-setting rule, as illustrated in Figure 4b: the reference wage increases, the lower threshold increases to $q^l(r_1) = r_1/[\lambda\eta] = q_0/\lambda$, and the upper threshold increases to $q^u(r_1) = r_1/\eta = q_0$.

If in the subsequent employment period the match productivity remains unchanged, so that $q_1 = q^u(r_1)$, optimal wage setting implies $\tilde{w}_1 = r_1 (= \tilde{w}_0)$. However notice that whilst the initial wage \tilde{w}_0 was positively reciprocated by the worker in period 0, after reference wage adaptation the worker has an updated sense of entitlement and now perceives this wage as fair, meaning that effort is merely normal $\tilde{e}_1 = \tilde{e}^n$. Hence, due to reference wage adaptation reciprocity is a temporary phenomenon. This implication of the model is consistent with the evidence reported by field surveys that the positive effects of a wage gift on morale and effort are believed to be only temporary by firms' managers (e.g. Campbell

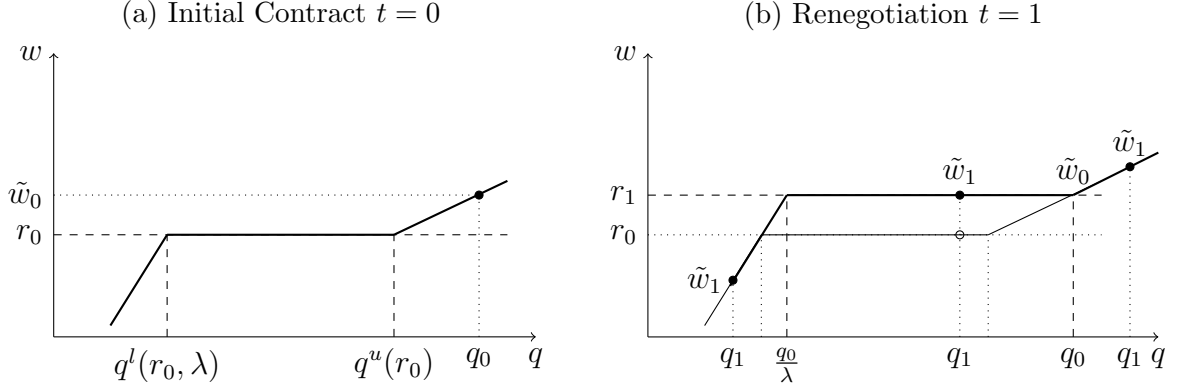


Figure 4:
Reference wage adaptation and downward wage rigidity.

and Kamlani (1997), Bewley (1999)). It also supports the interpretation according to which—in field experiments—positive reciprocity quickly disappears as workers get used to the wage they receive (see, for instance, Gneezy and List (2006), Mas (2006) and Cohn et al. (2014)). Evidence of such effort dynamics has also recently been documented in the laboratory experiment of Sliwka and Werner (2017), which also corroborates the hypothesis of reference wage adaptation.¹² We define this adjustment of effort over time as *dynamic ‘re-normalisation’ of effort*.

If the match productivity increases $q_1 > q_0$, the firm will instead find it optimal to increase the wage $\tilde{w}_1 = \eta q_1 > r_1 (= \tilde{w}_0)$, to benefit from the gift being reciprocated by supra-normal effort. On the other hand, if the match productivity decreases $q_1 < q_0$, whether the wage is adjusted downward depends on how large the negative shock is. As illustrated in Figure 4b, only if $q_1 < q_0/\lambda$ will the firm implement a wage cut $\tilde{w}_1 = \lambda \eta q_1 < r_1 (= \tilde{w}_0)$. As such, a fall in match productivity over time is not necessarily followed by a wage cut: the worker’s reference wage adaptation implies that, if the match productivity only moderately decreases $q_0/\lambda \leq q_1 < q_0$, the firm will optimally freeze the wage. The negative effect of what would now be perceived as an unfair wage cut, borne through negative reciprocity, will be larger than the benefit of paying a lower wage, hence the firm

¹²Sliwka and Werner (2017) conduct a laboratory experiment in which individuals work on a real-effort task and are paid different wage profiles which vary in the frequency and size of wage increases. They find that the positive effect on effort of a wage increase only lasts one period and that in the following periods, absent subsequent increases in the wage, working performance converges back towards the level associated with a constant wage. Interestingly, the field experiment by Kube et al. (2013) also indicates that negative reciprocity is more persistent than positive reciprocity. In light of our theory this evidence suggests that reference wage adaptation, which drives the temporary nature of reciprocity, may also be asymmetric: workers adapt more rapidly to wage increases than to wage cuts (see Fongoni (2018a) for a preliminary exploration of this hypothesis).

will avoid inciting such negative reciprocity by keeping the wage equal to the worker's reference wage $\tilde{w}_1 = r_1 (= \tilde{w}_0)$. To draw a more direct link with the empirical literature on downward wage rigidity (e.g. Dickens et al. (2007)), Figure 5 illustrates the distribution of log-wage changes, implied by our model, to a log-normally distributed shock around q_0 in period 1, comparing the case of symmetric ($\lambda = 1$) vs. asymmetric reference-dependent reciprocity ($\lambda > 1$).¹³

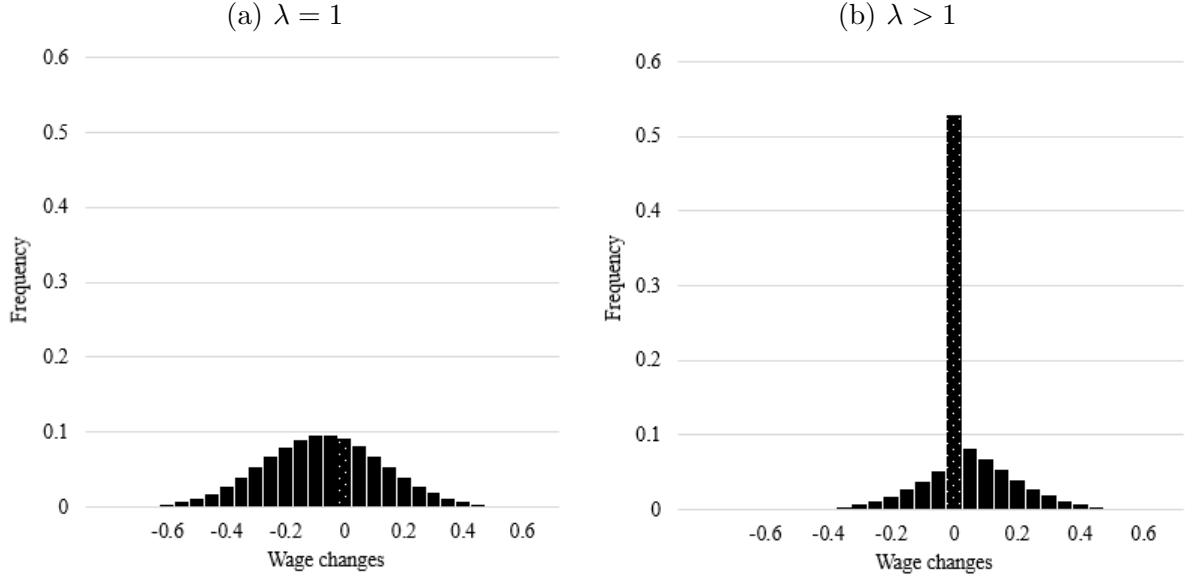


Figure 5:
Theoretical distributions of log-wage changes

Note: Simulated distributions of log-wage changes from the model based on the following assumptions and parameter values: the initial productivity and reference wage are $q_0 = 10$ and $r_0 = 1$; $\eta = 1$ and the loss aversion parameter is $\lambda = 1$ in Figure 5a and $\lambda = 1.8$ in Figure 5b (the experimental literature on loss aversion estimates $\lambda \in [1.43, 4.8]$; see Abdellaoui et al. (2007) for a review); $q_1 = q_0 + \varepsilon$, where the shock ε is log-normally distributed around 0 with variance 10, so that $\log q_1 \sim \mathcal{N}(q_0, 10)$; and normal effort is set as $\tilde{e}^n \equiv b = 10$ to ensure the employment relationship remains profitable even for large negative shocks. The number of simulations is 10,000.

Notice that the downward wage rigidity implied by our model is not a static result. By virtue of the worker's (one-period) adaptation to wage increases during periods of positive productivity growth, the firm will face the marginal trade-off between a wage cut and negative reciprocity at any subsequent employment period characterised by a fall in productivity. Downward wage rigidity may arise even at wage levels substantially higher than those with which the employment relationship had initially started. As such our theory formally demonstrates that downward wage rigidity is an inherent feature of the

¹³This simple simulation is inspired by the analogous exercise performed by Benjamin (2015). However note that contrary to his, but in line with empirically observed distributions of wage changes, our model does not generate a gap in the distribution of wage changes below zero.

employment contract in a dynamic environment, the key features that drive which are the worker's dynamic adaptation of the reference wage ($r_t = \tilde{w}_{t-1}$), and the relatively large cost to the firm of negative reciprocity that derives from loss aversion ($\lambda > 1$). In fact, the combination of these two features—more generally, an asymmetric adjustment cost around the past wage—is also necessary in other models of dynamic downward wage rigidity, such as, Elsby (2009), Holden and Wulfsberg (2009), Benjamin (2015) and Kaur (2018).

3.2 The optimal employment contract

We now turn back to the general model and analyse the properties of a forward-looking firm's wage setting and hiring decision in light of the behavioural mechanisms just discussed, which, whilst considered in the context of a parameterised model for illustration, apply equally to the general model.

We denote by $J_t(r_t, q_t)$ the firm's value function of the employment relationship in period $t = \{0, 1\}$. The functional equation corresponding to the firm's sequence problem in (9) can therefore be written as:

$$\begin{aligned} J_0(r_0, q_0) &= \max_{w_0} \{ \pi(w_0; q_0, \tilde{e}(w_0, r_0, \lambda)) + \delta \mathbb{E}_0[J_1(w_0, q_1)|q_0] \}, \\ \text{where } J_1(r_1, q_1) &= \max_{w_1} \pi(w_1; q_1, \tilde{e}(w_1, r_1, \lambda)). \end{aligned} \quad (10)$$

Due to reference wage adaptation $r_1 = w_0$, the expected continuation value of the employment relationship $\mathbb{E}_0[J_1(w_0, q_1)|q_0]$ now also depends on the initial wage. Recognising that the lay-off reservation productivity in period 1 may fall anywhere in the support of the distribution of match productivity, to ease notational burden we make the assumption that the parameters of the model are such that $\underline{q}(w_0, \lambda) < q^l(w_0, \lambda)$. Hence, the expected continuation value of the employment relationship can be expressed as

$$\begin{aligned} \mathbb{E}_0[J_1(w_0, q_1)|q_0] &= \int_{\underline{q}(w_0, \lambda)}^{q^l(w_0, \lambda)} J_1(w_0, q_1)^- dF(q_1|q_0) \\ &\quad + \int_{q^l(w_0, \lambda)}^{q^u(w_0)} J_1(w_0, q_1)^= dF(q_1|q_0) + \int_{q^u(w_0)}^{\infty} J_1(w_0, q_1)^+ dF(q_1|q_0), \end{aligned} \quad (11)$$

where $J_1(w_0, q_1)^{-;=;+}$ represents the continuation value of the employment relationship

when $w_1 < w_0$; $w_1 = w_0$; $w_1 > w_0$, in which effort is given by $\tilde{e}(w_1, w_0, \lambda)^-; \tilde{e}^n; \tilde{e}(w_1, w_0)^+$. This expression highlights that the firm faces different realisations of future profit when setting the initial wage w_0 , depending on whether the subsequent match productivity q_1 is below, within or above the range of rigidity $[q^l(w_0, \lambda), q^u(w_0)]$. Attentive observation of equation (11) allows us to infer two important insights: when setting the wage in the initial employment period the firm influences both the expected continuation value of the employment relationship in each of the three scenarios and the range of match productivity over which these scenarios occur.

Define the marginal effect of a wage increase in period 0 on the expected future profit in period 1 as

$$\Phi(w_0, \lambda) \equiv \frac{\partial}{\partial r_1} \int_{\underline{q}(w_0, \lambda)}^{\infty} J_1(w_0, q_1) dF(q_1|q_0).$$

We demonstrate in the following proposition that higher initial wages are always detrimental to expected future profit.

Proposition 1. *For all $\lambda \geq 1$, a higher initial wage reduces the expected continuation value of the employment relationship:*

$$\Phi(w_0, \lambda) = \int_{\underline{q}(w_0, \lambda)}^{q^l(w_0, \lambda)} y_e \tilde{e}_r(\tilde{w}_1, w_0, \lambda)^- dF - \int_{q^l(w_0, \lambda)}^{q^u(w_0)} s'(w_0) dF + \int_{q^u(w_0)}^{\infty} y_e \tilde{e}_r(\tilde{w}_1, w_0)^+ dF < 0.$$

When setting the initial wage in a dynamic environment a forward-looking firm will therefore account for an additional expected future *cost*: a higher initial wage increases the worker's reference wage in the subsequent renegotiation, which negatively influences effort and the value of the employment relationship to the firm.¹⁴ A marginal increase in the initial wage lowers this expected value because if the firm subsequently wants to cut the wage then the effect of negative reciprocity is greater; if it wishes to freeze the wage then the wage paid is simply higher; and if it wants to increase the wage then the effect of positive reciprocity is lower.

¹⁴While this prediction may seem obvious in the context of our model, notice that in models in which the worker's effort also depends on the absolute wage level (e.g. in the spirit of Akerlof (1982)), a higher initial wage will also generate an additional expected benefit in terms of higher effort in the future, in contrast to the result established in Proposition 1.

Next, define the marginal effect of a wage increase on the current profit in period 0 as

$$\Psi(w_0; q_0, r_0, \lambda) \equiv y_e(q, \tilde{e}(w_0, r_0, \lambda))\tilde{e}_w(w_0, r_0, \lambda) - s'(w_0).$$

So long as $w_0 \neq r_0$ the necessary first-order condition that characterises the solution to the firm's problem in (10) is

$$\Psi(w_0; q_0, r_0, \lambda) - \delta|\Phi(w_0, \lambda)| = 0. \quad (12)$$

The optimal hiring wage will be set to balance the inter-temporal tradeoff between the net marginal value in the initial period of a higher wage with the expected discounted marginal cost that stems from adaptation to this wage in the subsequent period.

As we note in the proof of the following theorem, for this condition to be also sufficient to characterise a maximum, it is required that the firm's value function $J_0(r_0, q_0)$ is concave in w_0 . This is not straightforward to prove since the sign of the derivative of $\Phi(w_0, \lambda)$ with respect to w_0 remains undetermined: $\Phi_{w_0} \leq 0$ (see Appendix for details).¹⁵ Nevertheless, note that the concavity of the instantaneous profit function established in Theorem 2 implies that $J_0(r_0, q_0)$ will also be concave if the firm is sufficiently impatient.

To proceed with the analysis we make the following assumption:

D3. The firm's discount factor δ is such that

$$\Psi_w(w_0; q_0, r_0, \lambda) - \delta|\Phi_{w_0}(w_0, \lambda)| < 0.$$

This assumption essentially implies that the firm will always put a larger weight on the 'current direct effect' of a change in the wage on marginal profit (captured by Ψ_w) than on the 'expected discounted future indirect effect' that results from the initial wage becoming the reference wage (captured by $\delta\Phi_{w_0}$).

Let $\hat{q}(r_0, \lambda, \delta)$ be the firm's reservation productivity that governs hiring, implicitly de-

¹⁵This technicality was also an issue for the characterisation of the optimal wage policy in the model of Elsby (2009), who considered an infinite-horizon environment. However, note that while Elsby (2009) had to resort to numerical simulations to prove concavity, since we are interested in the analytical characterisation of the optimal employment contract we will impose a simplifying assumption on the firm's discount factor instead.

defined by $J_0(r_0, q_0) = 0$. In a similar vein to our approach when considering a single employment period (i.e. Section 2), we can derive the optimal hiring wage of a forward-looking firm, the properties of which are presented in the following theorem.

Theorem 3. *The optimal hiring wage is given by*

$$\tilde{w}_0 = \hat{w}(r_0, q_0, \lambda, \delta) = \begin{cases} \hat{w}(r_0, q_0, \lambda, \delta)^+ & \text{if } q_0 > \hat{q}^u(r_0, \lambda, \delta) \\ r_0 & \text{if } \hat{q}^l(r_0, \lambda, \delta) \leq q_0 \leq \hat{q}^u(r_0, \lambda, \delta) \\ \hat{w}(r_0, q_0, \lambda, \delta)^- & \text{if } q_0 < \hat{q}^l(r_0, \lambda, \delta), \end{cases}$$

where $\hat{q}^l(r_0, \lambda, \delta)$ and $\hat{q}^u(r_0, \lambda, \delta)$ are respectively characterised by the value of q_0 such that

$$\lim_{\epsilon \rightarrow 0} \Psi(r_0 - \epsilon; q_0, r_0, \lambda) - \delta |\Phi(r_0 - \epsilon, \lambda)| = 0; \text{ and}$$

$$\lim_{\epsilon \rightarrow 0} \Psi(r_0 + \epsilon; q_0, r_0, \lambda) - \delta |\Phi(r_0 + \epsilon, \lambda)| = 0;$$

(all singletons), so long as $q_0 \geq \underline{\hat{q}}(r_0, \lambda, \delta)$.

The optimal wage $\hat{w}(r_0, q_0, \lambda, \delta)^{+(-)}$ is implicitly defined by (12). Moreover:

- a) $\hat{w}_q(r_0, q_0, \lambda, \delta) > 0$ for all $q_0 \in [\underline{\hat{q}}(r_0, \lambda, \delta), \infty) \setminus [\hat{q}^l(r_0, \lambda, \delta), \hat{q}^u(r_0, \lambda, \delta)]$; and
- b) $w_r(r_0, q_0, \lambda, \delta) > 0$ for all $q_0 \in [\underline{\hat{q}}(r_0, \lambda, \delta), \infty)$.

Model implications: downward wage rigidity and wage compression

Since a forward-looking firm perceives an additional marginal cost of raising the current wage (as we established in Proposition 1), it will have an incentive to hire the worker at a lower wage, relative to that of a myopic firm in an otherwise identical employment relationship. This insight, which we will refer to as ‘wage compression’, was first analysed by Elsby (2009), who attributes the incentive to compress the wage entirely to a firm’s anticipation of future downward wage rigidities.¹⁶ Our model identifies that it is the worker’s adaptation and re-normalisation of effort—and not downward wage rigidity *per se*—that is the main driver of wage compression. This is formally established by the following proposition.

¹⁶Also note that a similar wage compression effect is present in the DSGE model of Benigno and Ricci (2011) in which, however, downward wage rigidity is imposed as a purely exogenous constraint on the household optimisation problem (see the discussion at p.1444–1446).

Proposition 2. *For any $\lambda \geq 1$, a forward-looking firm will set a lower initial wage than a myopic firm:*

$$\hat{w}(r_0, q_0, \lambda, \delta) \leq \tilde{w}(r_0, q_0, \lambda),$$

with a strict inequality whenever $w_0 \neq r_0$.

The optimal wage (when not equal to the reference wage), is determined by the optimality condition in (12). While a myopic firm only considers the current net marginal value of a higher initial wage, a forward-looking firm also considers the expected discounted marginal cost of an increase in the hiring wage, which is positive for *all* workers, both those that are loss averse and those that are not, since

$$|\Phi(w_0, 1)| = \left| \int_{\underline{q}(w_0, 1)}^{\infty} y_e \tilde{e}_r(\tilde{w}_1, w_0, 1) dF \right| > 0.$$

This is due to reference wage adaptation and the dynamic re-normalisation of effort: a higher initial wage that is positively reciprocated in the first employment period would translate into a higher reference wage in the subsequent period, which in turn would reduce, in expectation, the worker's extent of reciprocity in the future—reciprocity in our model is a temporary phenomenon. As such, even in the absence of downward wage rigidity, a forward-looking firm has an incentive to compress the hiring wage.

Nevertheless, does the expectation of downward wage rigidity with a worker for whom $\lambda > 1$ reinforce or temper a firm's wage compression incentive? The answer to this question is particularly important in understanding whether the expectation of downward wage rigidity (for instance, of workers hired during recessionary episodes) would also imply a more compressed wage growth throughout the employment relationship. To provide an answer we need to analyse the effect of loss aversion λ on the optimal hiring wage \hat{w}_0 .¹⁷ Whenever $\tilde{w}_t < r_t$ for all $t = \{0, 1\}$, a firm employing a more loss averse worker has a stronger incentive to reduce the gap between the unfair wage paid and the reference wage, to attenuate the stronger effect of negative reciprocity (as we established in Section 2). In the initial employment period where the reference wage is exogenous, this puts upward

¹⁷Our analysis concerns the hiring wage since we consider a simple two-period environment. If we were to extend the time horizon, the results derived hereafter would in fact apply to any subsequent wage payment preceding the last employment period.

pressure on the hiring wage. We call this the *current direct effect* of loss aversion, denoted by $\Psi_\lambda > 0$. However, due to reference wage adaptation in the subsequent employment period: on one hand A) a greater extent of loss aversion puts downward pressure on the hiring wage, since by setting a lower wage—that will translate into a lower reference wage—the firm can reduce the magnitude of the expected negative reciprocity; but on the other hand B) since the firm will also be less willing to retain a worker who exhibits a stronger incidence of negative reciprocity (i.e., a more loss averse worker), the probability of the firm having to enact a wage cut is also lower, partially offsetting the greater expected cost of doing so.¹⁸ We define these as the *expected indirect effects* of loss aversion, given by¹⁹

$$\Phi_\lambda = \int_{\underline{q}}^{q^l} \underbrace{[y_{ee}\tilde{e}_\lambda\tilde{e}_r + y_e\tilde{e}_{r\lambda}]}_{A<0} dF - \underbrace{\underline{q}_\lambda y_e \tilde{e}_r f(\underline{q}|q_0)}_{B>0} \lesseqgtr 0. \quad (13)$$

The relative importance of these effects determines the overall incidence of loss aversion on the optimal wage contract.

Proposition 3. *The effect of λ on the hiring wage depends on the following conditions:*

- a) if $\Phi_\lambda < 0$ and $q_0 \geq \hat{q}^l(r_0, \lambda, \delta)$, then $\hat{w}_\lambda(r_0, q_0, \lambda, \delta) < 0$;
- b) if $\Phi_\lambda < 0$ and $q_0 < \hat{q}^l(r_0, \lambda, \delta)$, then $\hat{w}_\lambda(r_0, q_0, \lambda, \delta) < 0 \Leftrightarrow \Psi_\lambda < \delta|\Phi_\lambda|$;
- c) if $\Phi_\lambda \geq 0$, then $\hat{w}_\lambda(r_0, q_0, \lambda, \delta) \geq 0$.

Proposition 3 establishes that the effect of λ on the optimal hiring wage \hat{w}_0 is ambiguous, which is a very natural conclusion given the effects at play in our model. Suppose, for the initial part of this discussion, that the probability of being in a situation to enact costly wage cuts in the future is sufficiently high so that $\Phi_\lambda < 0$ (i.e. the lay-off reservation productivity $\underline{q}(w_0, \lambda)$ does not increase too much with λ , which implies that effect (B)

¹⁸This latter effect was established in Theorem 2. As such, our model predicts that in the presence of loss averse workers, there will be both downward wage rigidity and layoffs in periods in which productivity declines. Note, however, that if we were to simulate the model the full extent of downward wage rigidity may not be observed in the distribution of wage changes following a large negative shock even if it is a salient feature of the employment contract: firms that are critically constrained by downward wage rigidity will lay off the least productive workers. This is consistent with the recent work by Kurmann and McEntarfer (2017) who find that in the U.S. during the Great Recession wage freezes have been less popular, which they attribute not to a lack of downward wage rigidity but to the least productive workers being laid off, and to the most productive workers receiving wage cuts.

¹⁹For a detailed derivation of this expression, see the Proof of Proposition 3 in the Appendix.

identified above is sufficiently small). Now consider a worker that is hired at a relatively high match productivity $q_0 \geq \hat{q}^l(r_0, \lambda, \delta)$, i.e. case a) (hence $\hat{w}_0 \geq r_0$ so $\Psi_\lambda = 0$, i.e., there is no current direct effect of loss aversion). This scenario was also the one analysed in Section 3.1. In this particular case the expectation of stronger negative reciprocity and downward wage rigidity would unambiguously lead to wage compression, since the firm optimally sets a lower wage to keep the worker's wage entitlement low and reduce the extent of the expected negative reciprocity if $\tilde{w}_1 < \tilde{w}_0$ in the subsequent employment period.

Next consider a worker that is hired at a relatively low match productivity $q_0 < \hat{q}^l(r_0, \lambda, \delta)$, i.e., case b) (hence $\hat{w}_0 < r_0$ and $\Psi_\lambda > 0$). In this case, a higher λ would lead to wage compression only if the firm's incentive to set a higher wage to attenuate negative reciprocity in period 0 is dominated by the incentive to set a lower wage to keep the worker's reference wage low, and to reduce their negative reciprocity response in the event of a future (unfair) wage cut in period 1. However, notice that these conclusions are subject to presuming that $\Phi_\lambda < 0$. If, as in case c), $\Phi_\lambda \geq 0$, there is an additional marginal consideration that partially offsets the wage compression incentive: a greater probability of layoff following a large negative shock implies that the probability of the firm having to incur the cost of negative reciprocity in period 1 is also lower. As such we conclude that the expectation of downward wage rigidity does not necessarily lead to wage compression.²⁰ Our model highlights that the incentive for wage compression is driven by a worker's reference wage adaptation and dynamic re-normalisation of effort; and that wage rigidity may either strengthen or dampen this incentive.

²⁰Someone may argue that in a long-term employment relationship the initial effect of negative reciprocity will not be so important (i.e. $\Psi_\lambda < \delta|\Phi_\lambda|$ as $t \rightarrow \infty$) so that firms will always compress hiring wages. However, this statement is not necessarily true for two reasons. First, the expected future indirect effects of loss aversion are discounted, so it is still possible that the current direct effect dominates (i.e. $\Psi_\lambda > \delta|\Phi_\lambda|$ for sufficiently impatient firms). Second, in our model the expectation of downward wage rigidity also reduces the expected duration of the employment relationship (by increasing the firm's future lay-off reservation productivity), since firms would optimally layoff workers rather than implementing costly wage cuts. This latter effect (i.e. effect (B) of equation (13)) reduces the range of negative shocks over which a firm would experience negative reciprocity due to a wage cut, and therefore reduces the related costs of doing so. Finally notice that the models of Elsby (2009) and Benigno and Ricci (2011), which consider an infinite horizon, do not feature a lay-off reservation productivity that is endogenous to optimal wage setting and reciprocity, as we do in our model. That is why downward wage rigidity unambiguously implies wage compression in their models.

Model implications: loss aversion and hiring

We conclude our investigation of asymmetric reference-dependent reciprocity and wage rigidity on the nature of the employment contract by considering their effects on a firm's hiring decision. This can be done by considering the effect of loss aversion λ on the firm's *hiring* reservation productivity $\hat{q}(r_0, \lambda, \delta)$. There are two channels through which loss aversion could influence the value of the employment relationship to the firm. First, there is the direct negative effect on effort, which exacerbates a worker's negative reciprocity response in each $t = \{0, 1\}$. Second, there is the indirect effect that comes from our analysis related to Proposition 3: if the hiring wage is increasing in λ this provides a compounding negative effect on expected effort and a higher labour cost, which lowers profit; whilst if the initial wage is decreasing in λ —i.e. wage compression—there is a partially offsetting positive effect on expected effort and a lower labour cost, which increases profit. Nevertheless, we can show that if a firm is considering contracting with a more loss averse worker the reservation productivity determining hiring unambiguously increases, independently of how the hiring wage adjusts.

Proposition 4. *The firm's reservation productivity for hiring is increasing in λ :*

$$\hat{q}_\lambda(r_0, \lambda, \delta) > 0.$$

The mechanism behind the *neutrality* of this result to changes in the hiring wage lies in the firm's optimal wage setting: since the initial wage is set to balance the expected direct/indirect effects of loss aversion on output (effort) and labour cost (wage) at the margins to satisfy the first-order condition in (12), a higher degree of loss aversion negatively affects profit only through the stronger negative reciprocity response of the worker whenever $\tilde{w}_t < r_t$. We can therefore conclude that, independently of whether wage rigidity reinforces or tempers the incentive to compress hiring wages, the anticipation of stronger negative reciprocity and the expectation of downward wage rigidity unambiguously reduces a firm's incentive to hire.

This prediction is consistent with the empirical evidence reported in Kurmann and McEntarfer (2017), who find that firms that are more constrained by downward wage

rigidity employ on average higher productivity employees. Moreover, this result has implications for understanding the effects of wage rigidity on job creation. In the literature concerned with labour market fluctuations, much attention has been devoted to the effect of wage rigidity in the wages paid to *newly hired* workers on firms' job creation incentives (see for instance Pissarides (2009), the discussion in Elsby et al. (2015) and references therein). In contrast, Proposition 4 establishes that it is the anticipated negative reciprocity and the expected wage rigidity of *incumbents* that reduces a firm's incentive to hire, independently of the rigidity/flexibility of the hiring wage. As such, our analysis suggests that incorporating our model into a richer macroeconomic framework could potentially enhance the understanding of the effects of anticipated wage rigidities of *existing/incumbent* workers on job creation and unemployment.

4 Related Literature

The theory we develop in this paper reconciles within a single analytical framework several ideas and concepts that have been extensively discussed in relation to wage setting behaviour and the employment contract. As such, our paper is inevitably related with a large body of literature—spanning from the study of reciprocity in labour relations to wage dynamics.

4.1 Theories of reciprocity

Our model of asymmetric reference-dependent reciprocity is related to earlier theories of fairness and reciprocity that were developed to explain the growing body of experimental evidence against the self-interest hypothesis in economics. This literature can be divided between: i) pure outcome-based (or distributional preferences) reciprocity models, according to which individuals are only concerned about the distributional consequences of their actions (Fehr and Schmidt, 1999; Bolton and Ockenfels, 2000); ii) intentions-based reciprocity models, according to which it is a player's intention behind a certain action that matters for fairness judgements (Rabin, 1993; Dufwenberg and Kirchsteiger, 2004); and iii) a combination of both outcome- and intentions-based reciprocity (i.e. Falk and Fischbacher (2006)).

We consider a simple setting of reciprocity in which only one party to the transaction—the worker—has preferences for fairness captured in our morale function, and what they perceive as fair is given (by the reference outcome), rather than being constructed from what could otherwise have been chosen by the firm. Since the firm is assumed to know the worker’s perception of what is a fair wage, deviations from this are considered as being intended. The ‘kindness’ of the firm is captured by the wage paid in relation to the reference wage; whilst the ‘kindness’ of the worker is captured by their effort in relation to normal effort. The firm (who moves first) is assumed to be purely self-interested, but they nevertheless care about fairness because of the effect of the worker’s response to the wage paid on profit. Despite its simplicity, our model captures the essence of positive and negative reciprocity, and implies behaviour consistent with that documented in real labour markets.

4.2 Reciprocity in labour markets

Our paper is also related to the more recent theoretical literature that applies the concept of reciprocity to enhance our understanding of the employment relationship (microeconomic approaches) and labour market dynamics (macroeconomic approaches).

Microeconomic approaches. Benjamin (2015) extends the inequity-aversion model of Fehr and Schmidt (1999) to study a gift-exchange game between a self-interested profit-maximising firm and a ‘fair-minded’ worker, who chooses effort to equate their own ‘surplus payoff’ with that of the firm: whenever the firm’s profit increases (decreases) relative to a reference payoff, the worker will optimally decrease (increase) effort to re-balance the difference in payoffs. Particularly related to our derivation of asymmetric reciprocity and effort dynamics is the work of Sliwka and Werner (2017). In the spirit of Cox et al. (2007) they develop a dynamic model of reciprocity in which—depending on their ‘emotional state’—the worker also cares about the firm’s payoff. The reference wage, which affects the emotional state attached to a wage offer, follows a dynamic adaptation rule (equivalent to our Assumption D2). They test their model in the laboratory and find that the effect of wage gifts on effort is only temporary, as our model would also predict. However, Sliwka

and Werner (2017) do not analyse optimal wage setting and hiring behaviour.²¹ In a more recent paper Macera and te Velde (2018) develop a model of reference-dependent reciprocity and analyse the consequences of surprising *versus* anticipated wage gifts. They find that fully surprising gifts (when the reference wage is fixed) lead to higher effort than fully anticipated gifts (when the reference wage is determined by rational expectations à la Kőszegi and Rabin (2006)). Our application of reciprocity is fundamentally different from the ones just described which, essentially, are all based on the worker’s consideration of the firm’s payoff. Despite this leading to a set of predictions that are different from our model (and also hard to reconcile with existing evidence),²² we believe our model to be more plausible, since—as also argued by Dufwenberg and Kirchsteiger (2000, p.1071)—informational problems characterising actual labour relations can make it hard for workers to compare their payoff with that of their employer.²³ Finally notice that while these models rely on specific functional forms (and in the case of Benjamin (2015), also on numerical parameter values), our model of worker’s reciprocity is simpler, more general, and derives predictions with respect to optimal effort—namely, the asymmetric and temporary nature of reciprocity—by imposing relatively modest assumptions on the worker’s preferences.

Macroeconomic approaches. In a series of papers, Danthine and Kurmann (2006, 2007, 2010) assume that workers’ preferences exhibit reciprocity à la Rabin (1993) and analyse the macroeconomic implications for wage and unemployment dynamics. A distinctive feature of their framework is that the reference wage is also influenced by the firm’s ability to pay (the ‘internal reference’). In equilibrium, the firm’s gift is always positive and workers exert supra-normal effort. However, Danthine and Kurmann do not capture negative reciprocity in the form of sub-normal effort, and their model generates wage rigidity only under certain assumptions about the nature of shocks, workers’ reference wages and

²¹Despite being based on different assumptions regarding the worker’s preferences and reciprocity motives, our model and the one of Sliwka and Werner (2017) yield a similar theoretical prediction concerning the temporary nature of positive reciprocity subject to reference wage adaptation. In this respect we would like to acknowledge that the first version of our paper appeared in July 2015 (as a SIRE Discussion Paper SIRE-DP-2015-57). Therefore, we believe our framework and the one of Sliwka and Werner (2017) were likely to have been developed independently.

²²For instance: in Benjamin (2015) effort increases when the firm freezes the wage, and when the negative shock is so large that the firm optimally cuts the wage (see p. 201-202); in Sliwka and Werner (2017) the assumption on the worker’s preferences would also imply that an exogenous increase in the firm’s payoff will induce the worker to unconditionally exert higher effort; and in Macera and te Velde (2018) reciprocity is asymmetric only for small changes of the wage relative to the reference wage. These predictions seem hard to reconcile with existing evidence on reciprocity in labour markets.

²³See also the discussion in Kube et al. (2013, p. 12).

the functional form of the workers' gift. In a more recent contribution Eliaz and Spiegler (2014) qualitatively analyse the role of reference dependence, contractual incompleteness and negative reciprocity—subsumed under a reduced-form reference-dependent production function—for the volatility of unemployment. They capture an extreme form of negative reciprocity: a wage cut below the reference wage, no matter how small, will induce a worker to exert zero discretionary effort, leaving the adverse effect on output to be randomly determined by a parameter that represents the incompleteness of the labour contract.²⁴ In contrast, our model identifies the severity of the adverse effect of a wage cut on output by a combination of a worker's degree of loss aversion (which determines the strength of negative reciprocity) and the size of the wage cut.

4.3 Loss aversion and wage setting in labour markets

A key implication of our model is that wage cuts relative to the reference wage have a stronger, negative impact on workers' morale and effort than equivalent-sized wage gifts, which stems from loss aversion in the worker's evaluation of the wage relative to a reference wage.

Loss aversion in the context of the labour market was first analysed by Bhaskar (1990), who considered its implications for the determination of equilibrium unemployment in a static monopoly-union model in which workers care about relative wage comparisons. In a similar vein McDonald and Sibly (2001) incorporate loss aversion into an insider-outsider model and focus on the implications for monetary policy. They show that under the assumption of loss aversion, reference wage adaptation, and staggered wage contracts, unanticipated monetary shocks can have permanent real effects on output and unemployment. More recently Ahrens et al. (2015) analyse the quantitative implications of reference dependence and loss aversion in a monopsony model of the labour market in which workers decide on how much employment (hours) to supply. According to their numerical simula-

²⁴As they show in an appendix, Eliaz and Spiegler (2014) can derive their reduced-form production function from a simple model that yields an expression for discretionary effort: if a worker is paid a wage at least equal to their reference wage, the worker is *assumed* to exert effort normalised to unity (see p. 195); if they are paid an unfair wage below their reference wage, effort is zero. Since the relative importance of discretionary effort in a firm's output is given by the extent to which the labour contract is incomplete, this also determines the random fraction of output that is destroyed when a worker is paid below their reference wage. As such, Eliaz and Spiegler's (2014) model suggests that the more a contract is incomplete, the greater is the adverse effect of wage cuts on output.

tion, loss aversion implies that whether wages are adjusted crucially depends on the size and sign of labour demand shocks: for small shocks around the steady state, wages are completely rigid (both upward and downward); for large shocks, wages are either downward rigid, or relatively downward sluggish, or upward flexible.

Finally, note that the models by Benjamin (2015), Sliwka and Werner (2017) and Macera and te Velde (2018), discussed above, also incorporate loss aversion into the worker’s preferences.²⁵ However, only Benjamin (2015) studies downward wage rigidity (see discussion below), while Sliwka and Werner (2017) and Macera and te Velde (2018) do not analyse the implications of loss aversion for dynamic downward wage rigidity, wage compression and hiring, which is the main contribution of our paper.

4.4 The reference ‘fair’ wage

Another important element of our model is the adaptation of the worker’s reference wage to the past wage. This assumption is consistent with a large body of evidence documented in the labour market literature on reference wage formation as well as by other behavioural science sub-disciplines concerned with reference point formation.

The first piece of evidence supporting the idea that past wage contracts serve as a reference for fairness judgements in the labour market comes from the seminal experiment of Kahneman et al. (1986). This finding has been rapidly coupled with the psychological notion of adaptation, or habituation, popular in social psychology (see Kahneman and Thaler (1991) and Baucells and Sarin (2010) for a review of this early literature). Adaptation to past wage contracts is also supported by several anthropological studies (see the survey of Bewley (2007)). In every firm interviewed, regardless of the country or industry of origin, compensation managers believe that past wage contracts are important determinants of workers’ wage entitlements, and that workers rapidly adapt to what they are paid. In the context of experimental studies, indirect evidence in support of this hypothesis comes from the field experiments of Gneezy and List (2006) and Mas (2006), and the laboratory experiments of Clark et al. (2010), Gächter and Thöni (2010) and Koch (2017) among others.

²⁵However notice that in contrast to our framework, in the model of Benjamin (2015) workers are loss averse in their wage *and* effort domains, and in Macera and te Velde (2018) workers also exhibit loss aversion when the firm’s profit decreases relative to a reference profit level.

Direct evidence of reference wage adaptation and its effect on the dynamics of positive reciprocity has also been documented by the field experiment of Chemin and Kurmann (2014) and the laboratory experiment of Sliwka and Werner (2017). Moreover, the idea that ex-ante contracts serve as entitlements for future renegotiations was advanced by Hart and Moore (2008) and further explored in Herweg and Schmidt (2012) in the literature on incomplete contracts. The laboratory experiments of Fehr et al. (2011, 2014), Bartling and Schmidt (2015) and Herz and Taubinsky (2018) provide strong support for this hypothesis.

Taken together this evidence identifies the past wage as one of the most plausible candidates for a worker’s reference wage, pointing to the existence of a dynamic adaptation of wage entitlements over time. That said, there are other plausible determinants of the reference wage that may differ depending on whether the worker is a new hire (in period 0) or an incumbent (in period 1). For instance a *newly hired* worker’s reference wage r_0 could be influenced by the state of the labour market (as in the efficiency wage tradition, e.g. Akerlof (1982) and Summers (1988)), the most recent wage contract paid in the *previous* employment relationship (as considered in Koenig et al. (2014)), or the wage of incumbent workers employed by the same firm (as the “equal treatment” hypothesis of Snell and Thomas (2010) would suggest). On the other hand an *existing/incumbent* worker’s reference wage r_1 might be influenced by the wage of his peers outside the firm (as in Keynes (1936), Bhaskar (1990) or Driscoll and Holden (2004)), by expectations (as in Eliaz and Spiegler (2014) and Macera and te Velde (2018)) or by the firm’s ability to pay (as in Danthine and Kurmann (2007)). We take a simple approach in this article that nevertheless generates predictions in line with stylised facts, but a very interesting direction for future research is to gain further understanding of what influences the reference wage in a labour market setting, and determine whether or not adaptation is symmetric for wage increases and reductions.

4.5 Models of dynamic downward wage rigidity

A key contribution of our paper is to provide a transparent, and plausible, microfoundation for downward wage rigidity in a dynamic environment. The crucial role that wage rigidity plays in the understanding of business cycles and labour market fluctuations has

contributed to the development of a rich set of theoretical models of wage rigidity/stickiness throughout the history of macroeconomics.²⁶ While a complete review of this literature is beyond the scope of the present paper, we focus the discussion that follows on theoretical models of downward wage rigidity that: i) can reproduce, or are consistent with, observed empirical distributions of wage changes; ii) generate downward wage rigidity in a dynamic environment; and iii) are based on well-documented stylised facts on the behaviour of workers and firms engaged in employment relationships. We believe these features to be important for a microeconomic theory that aims to provide a plausible account of firms' wage setting behaviour and observed wage dynamics.

As also pointed out by Snell et al. (2018, p. 8), there are few models of wage setting that can generate wage dynamics consistent with i) and ii) above. We highlighted in Section 3.1 that a necessary ingredient for this is an asymmetric adjustment cost around the past wage: the marginal cost of cutting the wage (in terms of effort) needs to be larger than the marginal benefit of increasing it. Indeed, for this asymmetry to be always present in a dynamic environment it is also necessary for it to be centered around the past wage. In our model these features are captured by the worker's asymmetric reference-dependent reciprocity combined with the dynamic adaptation of the reference wage to past wage contracts, consistent with point iii) above.

Models of wage setting that generate dynamic downward wage rigidity are those of Elsby (2009), Holden and Wulfsberg (2009), Eliaz and Spiegler (2014) and Kaur (2018). However, while being informed by the large body of literature on loss aversion and reciprocity, they all consider reduced-form functions to characterise the behaviour of workers: both Elsby (2009) and Kaur (2018) consider a reduced-form effort function with a kink at the past wage; Holden and Wulfsberg (2009) consider a utility function that exhibits a kink at the point where the wage is equal to the past wage; and Eliaz and Spiegler (2014) consider a reduced-form production function in which output falls disproportionately if wages are set below the worker's *lagged*-expectations about the wage. Dynamic downward wage rigidity

²⁶To avoid confusion, we define wage rigidity as the *acyclical* behaviour of wages, i.e. when wages do not adjust to shocks (downward/upward or both); while wage stickiness is defined as a *less than proportional cyclical* of wages with respect to shocks. In the literature these terms have been used interchangeably, sometimes referring to the latter and sometimes to the former. An excellent and comprehensive survey of earlier theories of wage rigidity/stickiness can be found in Bewley (1999), while more recent surveys are present in, for instance, Babecký et al. (2010) and Rogerson and Shimer (2011).

is also a feature of the inequity-aversion model of Benjamin (2015) once loss aversion—in the domain of wages and effort—is introduced on the worker’s side. However, despite this implying that inequity-aversion *per se* is not enough to explain downward wage rigidity in a dynamic environment, as we also discussed in footnotes 13 and 22 the model of Benjamin is not entirely consistent with points i) and iii) listed above.

The model presented here provides a micro foundation—based on loss aversion, asymmetric reciprocity and reference wage adaptation—that supports dynamic downward wage rigidity and is consistent with points i), ii) and iii).

5 Conclusion

Inspired by evidence from anthropological and experimental research on labour markets, in this paper we have advanced a microeconomic theory of asymmetric reciprocity and optimal wage setting based on contractual incompleteness, fairness, and reference dependence and loss aversion in the evaluation of wage contracts by workers. This approach allows us to rigorously formalise several aspects of wage and effort dynamics, and to study their implications for the nature of the employment contract within a plausible and tractable model. By establishing a clear link between assumptions and conclusions, our theory provides novel insights to explain the observed asymmetry and dynamics of workers’ reciprocity, and to identify the sources of downward wage rigidity, wage compression and hiring incentives.

We formally characterise a worker’s effort response to wage changes to be reference dependent, where positive and negative reciprocity are defined as relative deviations from normal effort, and loss aversion is identified as the psychological foundation for the stronger intensity of negative reciprocity. In addition, we show that the reference-dependent nature of reciprocity, combined with adaptation, leads to a dynamic ‘re-normalisation’ of effort throughout the employment relationship: reciprocity is a temporary phenomenon. This prediction is consistent with the recent experimental findings of Sliwka and Werner (2017) and with other evidence documented in the field (e.g. Mas (2006) and Gneezy and List (2006)).

By subsequently analysing the implications of our theory of asymmetric reference-dependent reciprocity for optimal wage setting in a two-period environment, we establish

that downward wage rigidity is an inherent feature of the employment contract; and we identify an asymmetric adjustment cost around the past wage to be a necessary ingredient for models of dynamic downward wage rigidity. In our model this mechanism is generated by the worker’s adaptation of the reference wage and the relatively large cost of negative reciprocity in response to wage cuts (that stems from loss aversion). As such we think of our model as a general and plausible micro-foundation for downward wage rigidity in a dynamic environment.

When analysing the consequences of these wage and effort dynamics for the optimal employment contract, we draw new conclusions about their implications for a forward-looking firm’s wage compression incentive (Elsby, 2009), and for the expected value of the employment relationship, which influences hiring decisions. We find that the primary behavioural mechanism that generates wage compression is the anticipation of the worker’s re-normalisation of effort due to adaptation, even absent downward wage rigidity. In a model in which layoffs are endogenous to optimal wage setting, the expectation of downward wage rigidity may not necessarily lead to wage compression. Nevertheless, independently of how the hiring wage adjusts, the anticipation of stronger negative reciprocity and the expectation of downward wage rigidity unambiguously reduces the expected value of the employment relationship, implying that a firm that expects to be constrained by downward wage rigidity in the future will hire less on average.

The framework developed in this paper lends itself as a tractable benchmark model for the analysis of reference-dependent reciprocity, adaptation and wage rigidity, and their effect on wage setting and hiring behaviour. We have identified two extensions. First, it will be interesting to analyse the model predictions under different specifications of a worker’s reference wage. We chose the past wage as the only determinant of an incumbent worker’s reference wage as it is the most corroborated hypothesis in the empirical literature. However, like every model based on reference dependence, predictions are sensitive to the choice of reference point and investigating if, and how, our conclusions might change is the natural next step. Second, we believe that exploring the insights of the model within a richer macroeconomic framework can shed new light on the effects of expected wage rigidity in long-term employment relationships for job creation and wage dynamics (see Fongoni

(2018b) for a first step into this direction). As discussed in Elsby et al. (2015) this aspect is not yet settled in the theory of labour market fluctuations, and has drawn particular attention in light of recent cross-country experiences following the Great Recession (Elsby et al., 2016). Therefore, a promising line of future research lies in developing and combining these two extensions.

Appendix: Additional Material

Deriving the worker's participation constraint

We want to derive an expression for the worker's reservation wage defining their participation constraint. In our model this is the wage below which a worker will optimally turn down a job offer and stay unemployed in period 0, or quit the job and move to unemployment in period 1. Denote by $W_t(w_t, r_t)$ the value of the employment relationship to the worker in period t and normalise the value of being unemployed to zero. The worker's reservation wage in each period can be defined as: $\underline{w}(r_t, q_t) = \{0, w_t : W_t(w_t, r_t) = 0\}$.

By noting that the worker's decision to turn down a job offer would be made after observing the optimal wage contract set by the firm, and considering that $r_1 = \tilde{w}_0$ due to adaptation, we can express the value of the two-period employment relationship to the worker as

$$W_0(\tilde{w}(r_0, q_0), r_0) = u_0(\tilde{e}(\tilde{w}(r_0, q_0), r_0, \lambda); \tilde{w}(r_0, q_0), r_0) + \gamma \int_{\underline{q}(\tilde{w}(r_0, q_0), \lambda)}^{\infty} W_1(\tilde{w}(\tilde{w}(r_0, q_0), q_1), \tilde{w}(r_0, q_0)) dF(q_1|q_0),$$

where γ is the worker's discount factor and the optimal wage has been denoted here by $\tilde{w}(r_t, q_t)$ (i.e. excluding the other time-invariant functional arguments to ease notation). This expression can then be used to derive the reservation wage $\underline{w}(r_t, q_t)$ for each $t = \{0, 1\}$, the effect of which is to add an additional threshold to the model.

However, note that the reservation wage in each period will be a function of exogenous variables only, i.e. the initial reference wage r_0 and the initial match productivity q_0 . Hence it would be straightforward to impose a condition on these such that the worker's

participation constraint is never binding. While doing so will not affect the results presented in this paper, the modelling of this condition might be relevant in a richer framework in which the worker's initial reference wage is endogenous, and, in particular, dependant on the state of the labour market.

Appendix: Proofs

Proof of Theorem 1. We suppose Assumptions W1-W3 hold throughout. To begin note that if the wage is too low, $w < \underline{w}(r, \lambda)$, defined such that $-d'(0) + n(\underline{w}(r, \lambda)|r) = 0$, the worker would want to choose negative effort but is constrained not to do so. Hence for $w \leq \underline{w}(r, \lambda)$ we define $\tilde{e}(w, r, \lambda) \equiv 0$, and so long as $w > \underline{w}(r, \lambda)$ optimal effort will be strictly positive, a sufficient condition for which is $|d'(0)| > \lambda \eta m(r)$, which we assume to be the case throughout. So long as $w > \underline{w}(r, \lambda)$ optimal effort is given by the inverse function $\tilde{e}(w, r, \lambda) = d'^{-1}(n(w|r))$ which exists since d' is strictly monotonic. Since d' is a continuous function and $n(w|r)$ varies continuously in w and r , $\tilde{e}(w, r, \lambda)$ will be a continuous function of w and r , but it will not be continuously differentiable everywhere as, recalling its definition from (4), $n(w|r)$ has a kink at $w = r$. For $w \neq r$ we can apply the inverse function theorem to give

$$\tilde{e}_w(w, r, \lambda) = \begin{cases} \frac{\eta m'(w)}{d''(e)} & \text{if } w > r \\ \frac{\lambda \eta m'(w)}{d''(e)} & \text{if } w < r \end{cases}$$

so $\tilde{e}_w(w, r, \lambda) > 0$ for all $w \neq r$, and it then follows that

$$\tilde{e}_{ww}(w, r, \lambda) = \begin{cases} \frac{\eta m''(w)}{d''(e)} & \text{if } w > r \\ \frac{\lambda \eta m''(w)}{d''(e)} & \text{if } w < r \end{cases}$$

so $\tilde{e}_{ww}(w, r, \lambda) < 0$ for all $w \neq r$.

By appealing to the definition of normal effort when $w = r$, we can deduce that

$$\begin{aligned}\lim_{\epsilon \rightarrow 0} \tilde{e}_w(r - \epsilon, r, \lambda)^- &= -\lim_{\epsilon \rightarrow 0} \frac{\lambda \eta m'(r - \epsilon)}{d''(\tilde{e}(r - \epsilon, r, \lambda)^-)} = -\frac{\lambda \eta m'(r)}{d''(\tilde{e}_n)} \\ &= -\lambda \lim_{\epsilon \rightarrow 0} \frac{\eta m'(r + \epsilon)}{d''(\tilde{e}(r + \epsilon, r)^+)} = \lambda \lim_{\epsilon \rightarrow 0} \tilde{e}_w(r + \epsilon, r)^+.\end{aligned}$$

Hence the effort function kinks to a flatter slope as the wage increases. Note that this, combined with the deduction that $\tilde{e}_{ww} < 0$ for all $w \neq r$, implies \tilde{e}_w is everywhere decreasing in w , i.e. the effort function is concave.

The inverse function theorem and the definition of $n(w|r)$ can then be used to deduce the remainder of the claims:

$$\tilde{e}_r(w, r, \lambda) = \begin{cases} -\frac{\eta m'(r)}{d''(e)} & \text{if } w > r \\ -\frac{\lambda \eta m'(r)}{d''(e)} & \text{if } w < r \end{cases}$$

so $\tilde{e}_r(w, r, \lambda) < 0$ for all $w \neq r$,

$$\tilde{e}_\lambda(w, r, \lambda) = \begin{cases} 0 & \text{if } w > r \\ \frac{\eta[m(w) - m(r)]}{d''(e)} < 0 & \text{if } w < r \end{cases}$$

and so for $w < r$

$$\tilde{e}_{w\lambda}(w, r, \lambda) = \frac{\eta m'(w)}{d''(e)} > 0.$$

□

Proof of Theorem 2. Throughout the proof we assume the worker's productivity and reference wage are such that $q \geq \underline{q}(r, \lambda)$ so the firm will be profitable if it hires the worker, and consider the properties of the threshold productivity at the end. We proceed by first stating some preliminaries, then considering the productivity thresholds, then demonstrating the nature of the optimal wage setting rule.

Preliminaries: To ease notational burden denote the left-hand side of the first-order condition (7), i.e. the marginal profit, by $\Psi(w; q, r, \lambda)$. First, note that under Assumption F2 and the results of Theorem 1, for $w \neq r$ we have that $\Psi_q(w; q, r, \lambda) = y_{qe} \tilde{e}_w > 0$; $\Psi_r(w; q, r, \lambda) = y_{ee} \tilde{e}_r \tilde{e}_w + y_e \tilde{e}_{wr} \geq 0$ (after noting that $\tilde{e}_{wr} = 0$); and $\Psi_w(w; q, r, \lambda) =$

$y_{ee}[\tilde{e}_w]^2 + y_e \tilde{e}_{ww} < 0$. In addition, $\Psi_\lambda = y_{ee} \tilde{e}_\lambda \tilde{e}_w + y_e \tilde{e}_{w\lambda}$ so $\Psi_\lambda > 0$ if $w < r$ and $\Psi_\lambda = 0$ if $w > r$. These results also allow us to deduce that if $\lambda > 1$, $\Psi(w; q, r, \lambda)$ jumps down at the reference wage, since

$$\begin{aligned}
\lim_{\epsilon \rightarrow 0} \Psi(r - \epsilon; q, r, \lambda) - \lim_{\epsilon \rightarrow 0} \Psi(r + \epsilon; q, r, \lambda) &= \\
&= y_e(q, \lim_{\epsilon \rightarrow 0} \tilde{e}(r - \epsilon, r, \lambda)^-) \lim_{\epsilon \rightarrow 0} \tilde{e}_w(r - \epsilon, r, \lambda)^- - s'(r - \epsilon) \\
&\quad - y_e(q, \lim_{\epsilon \rightarrow 0} \tilde{e}(r + \epsilon, r, \lambda)^+) \lim_{\epsilon \rightarrow 0} \tilde{e}_w(r + \epsilon, r, \lambda)^+ + s'(r + \epsilon) \\
&= y_e(q, \tilde{e}_n) [\lim_{\epsilon \rightarrow 0} \tilde{e}_w(r - \epsilon, r, \lambda)^- - \lim_{\epsilon \rightarrow 0} \tilde{e}_w(r + \epsilon, r, \lambda)^+] \\
&= y_e(q, \tilde{e}_n) \lim_{\epsilon \rightarrow 0} \tilde{e}_w(r + \epsilon, r, \lambda)^+ [\lambda - 1] \geq 0,
\end{aligned}$$

with a strict inequality if $\lambda > 1$. As such, $\Psi(w; q, r, \lambda)$ is everywhere decreasing in w , establishing concavity of the firm's profit function.

Productivity thresholds: The threshold $q^l(r, \lambda)$ identifies the critical match productivity below which the firm would want to set the wage below the reference wage, and $q^u(r)$ is the match productivity above which the firm would want to compensate the worker more than the reference wage. The former is the value of q below which profit is decreasing just below the reference wage (so concavity of profit and the fact that $\Psi_q > 0$ imply that for all q below this, profit will be maximised when $w < r$); the latter is the value of q above which profit is increasing just above the reference wage (so concavity of profit and $\Psi_q > 0$ imply that for all q exceeding this, profit will be maximised when $w > r$). Since $\Psi(w, 0, r, \lambda) < 0$ when $w > 0$ and $\Psi_q > 0$ there will be a unique value of each productivity threshold.

We now want to establish some properties of the thresholds. Implicit differentiation allows us to deduce that

$$\begin{aligned}
q_r^l(r, \lambda) &= - \lim_{\epsilon \rightarrow 0} \frac{\frac{d\Psi(r - \epsilon; q, r, \lambda)}{dr}}{\Psi_q(r - \epsilon; q, r, \lambda)} \text{ and} \\
q^{u'}(r) &= - \lim_{\epsilon \rightarrow 0} \frac{\frac{d\Psi(r + \epsilon; q, r, \lambda)}{dr}}{\Psi_q(r + \epsilon; q, r, \lambda)}.
\end{aligned}$$

Now,

$$\begin{aligned}
\frac{d\Psi(r \pm \epsilon; q, r, \lambda)}{dr} &= \Psi_w(r \pm \epsilon; q, r, \lambda) + \Psi_r(r \pm \epsilon; q, r, \lambda) \\
&= y_{ee}[\tilde{e}_w^\pm]^2 + y_e \tilde{e}_{ww}^\pm + y_{ee} \tilde{e}_r^\pm \tilde{e}_w^\pm \\
&= y_{ee} \tilde{e}_w^\pm [\tilde{e}_w^\pm + \tilde{e}_r^\pm] + y_e \tilde{e}_{ww}^\pm.
\end{aligned}$$

As $\epsilon \rightarrow 0$ we can infer that $\tilde{e}_w(r \pm \epsilon, r, \lambda)^\pm + \tilde{e}_r(r \pm \epsilon, r, \lambda)^\pm \rightarrow 0$ (refer to the expressions of these objects in the proof of Theorem 1), implying $\lim_{\epsilon \rightarrow 0} \frac{d\Psi(r \pm \epsilon; q, r, \lambda)}{dr} = y_e \tilde{e}_{ww}^\pm < 0$. This allows us to conclude that $q_r^l(r, \lambda) > 0$ and $q^{u'}(r) > 0$.

Turning next to investigate how the lower threshold depends on the degree of loss aversion, implicit differentiation gives

$$\begin{aligned}
q_\lambda^l(r, \lambda) &= -\lim_{\epsilon \rightarrow 0} \frac{\Psi_\lambda(r - \epsilon, q, r, \lambda)}{\Psi_q(r - \epsilon, q, r, \lambda)} \\
&= -\frac{y_{ee} \lim_{\epsilon \rightarrow 0} \tilde{e}_\lambda(r - \epsilon, r, \lambda)^- \lim_{\epsilon \rightarrow 0} \tilde{e}_w(r - \epsilon, r, \lambda)^- + y_e \lim_{\epsilon \rightarrow 0} \tilde{e}_{w\lambda}(r - \epsilon, r, \lambda)^-}{\lim_{\epsilon \rightarrow 0} \Psi_q(r - \epsilon; q, r, \lambda)} < 0
\end{aligned}$$

since we deduced in Theorem 1 that $\tilde{e}_\lambda^- < 0$ and $\tilde{e}_{w\lambda}^- > 0$.

The consideration of $\lim_{\epsilon \rightarrow 0} \Psi(r - \epsilon; q, r, \lambda) - \lim_{\epsilon \rightarrow 0} \Psi(r + \epsilon; q, r, \lambda)$ in the preliminaries allows us to conclude that when $\lambda = 1$ these two objects are equal. This, combined with the observation that $\lim_{\epsilon \rightarrow 0} \tilde{e}(r - \epsilon, r, \lambda)^- = \tilde{e}_n = \lim_{\epsilon \rightarrow 0} \tilde{e}(r + \epsilon, r)$ (from Theorem 1) permits the conclusion that $q^l(r, 1) = q^u(r)$. This, along with the fact that $q_\lambda^l(r, \lambda) < 0$, implies $q^l(r, \lambda) < q^u(r)$ for all $\lambda > 1$.

Optimal wage setting: We now turn to the optimal wage setting rule, which depends on the match productivity in relation to the productivity thresholds.

If $q \in [\underline{q}(r, \lambda), q^l(r, \lambda))$ then the definition of $q^l(r, \lambda)$ and fact that $\Psi_q > 0$ can be used to deduce that $\lim_{\epsilon \rightarrow 0} \Psi(r - \epsilon, q, r, \lambda) < 0$; since $\Psi(w; q, r, \lambda)$ is everywhere decreasing in w , the same is true for all $w \geq r$. As such, the optimising wage must satisfy $w < r$ and will therefore be the solution to

$$y_e(q, \tilde{e}(w, r, \lambda)^-) \tilde{e}_w(w, r, \lambda)^- - s'(w) \leq 0,$$

with equality if $w > \underline{w}(r, \lambda)$ (recall from the proof of Theorem 1 that this is the wage below

which effort takes the boundary value of zero). To account for the fact that the firm may pay the ‘lowest feasible wage’ for a range of match productivity, let $\check{q}(r, \lambda) = \max\{0, q : \Psi(\underline{w}(r, \lambda); q, r, \lambda) = 0\}$ (at $\check{q}(r, \lambda)$ the firm would want to pay $\underline{w}(r, \lambda)$ and since $\Psi_q > 0$ the same will be true for all $0 \leq q < \check{q}(r, \lambda)$). For all $\check{q}(r, \lambda) < q < q^l(r, \lambda)$ the optimal wage is given by the displayed first-order condition holding with equality, which is denoted by $\tilde{w}(r, q, \lambda)^-$. Implicit differentiation and our deductions in the preliminaries reveal

$$\begin{aligned}\tilde{w}_q(r, q, \lambda)^- &= -\frac{\Psi_q}{\Psi_w} > 0, \\ \tilde{w}_r(r, q, \lambda)^- &= -\frac{\Psi_r}{\Psi_w} \geq 0, \text{ and} \\ \tilde{w}_\lambda(r, q, \lambda)^- &= -\frac{\Psi_\lambda}{\Psi_w} > 0.\end{aligned}$$

If $q \in (q^u(r), \infty)$ then the definition of $q^u(r)$ and the fact that $\Psi_q > 0$ can be used to deduce that $\lim_{\epsilon \rightarrow 0} \Psi(r + \epsilon, q, r, \lambda) > 0$; since $\Psi(w; q, r, \lambda)$ is everywhere decreasing in w the same is true for all $w \leq r$ and, as such, the optimising wage must exceed r and will therefore satisfy

$$y_e(q, \tilde{e}(w, r)^+) \tilde{e}_w(w, r)^+ - s'(w) = 0.$$

Letting $\tilde{w}(q, r)^+$ denote the solution (which is independent of λ), implicit differentiation gives

$$\begin{aligned}\tilde{w}_q(r, q)^+ &> 0 \text{ and} \\ \tilde{w}_r(r, q)^+ &\geq 0.\end{aligned}$$

If $q \in [q^l(r, \lambda), q^u(r)]$ then the fact that $\Psi_q > 0$ can be used to deduce that $\lim_{\epsilon \rightarrow 0} \Psi(r - \epsilon, q, r, \lambda) \geq 0$ and $\lim_{\epsilon \rightarrow 0} \Psi(r + \epsilon, q, r, \lambda) \leq 0$. That $\Psi_w < 0$ for all $w \neq r$ then implies $\Psi(w; q, r, \lambda) > 0$ for all $w < r$ and $\Psi(w; q, r, \lambda) < 0$ for all $w > r$, implying profit is maximised if and only if $w = r$.

Finally, if $q < \underline{q}(r, \lambda)$ then the employment relationship ends. Implicit differentiation of the zero profit condition defining the reservation productivity allows us to deduce

that

$$\begin{aligned}\underline{q}_r(r, \lambda) &= -\frac{y_e[\tilde{e}_r + \tilde{e}_w \tilde{w}_r] - s' \tilde{w}_r}{y_q + y_e \tilde{e}_w \tilde{w}_q - s' \tilde{w}_q} \\ &= -\frac{\tilde{w}_r[y_e \tilde{e}_w - s'] + y_e \tilde{e}_r}{\tilde{w}_q[y_e \tilde{e}_w - s'] + y_q} > 0\end{aligned}$$

since $y_e \tilde{e}_w - s' = 0$ from the first-order condition, $y_q, y_e > 0$ by Assumption F2 and we found in Theorem 1 that $\tilde{e}_r < 0$. In addition,

$$\begin{aligned}\underline{q}_\lambda(r, \lambda) &= -\frac{y_e[\tilde{e}_\lambda + \tilde{e}_w \tilde{w}_\lambda] - s' \tilde{w}_\lambda}{y_q + y_e \tilde{e}_w \tilde{w}_q - s' \tilde{w}_q} \\ &= -\frac{\tilde{w}_\lambda[y_e \tilde{e}_w - s'] + y_e \tilde{e}_\lambda}{\tilde{w}_q[y_e \tilde{e}_w - s'] + y_q}.\end{aligned}$$

Again $y_e \tilde{e}_w - s' = 0$ and $y_q, y_e > 0$, and we found in Theorem 1 that when $w > r$ \tilde{e} is independent of λ , but when $w < r$, $\tilde{e}_\lambda < 0$. As such, if $\tilde{w}(r, \underline{q}(r, \lambda), \lambda) > r$ then $\underline{q}_\lambda(r, \lambda) = 0$ but if $\tilde{w}(r, \underline{q}(r, \lambda), \lambda) < r$, $\underline{q}_\lambda(r, \lambda) > 0$. \square

Proof of Proposition 1. Recall that $J_1(w_0, q_1)^{-;+}$ represents the continuation value of the employment relationship when $w_1 < w_0$; $w_1 = w_0$; $w_1 > w_0$, in which effort is given by $\tilde{e}(w_1, w_0, \lambda)^-; \tilde{e}^n; \tilde{e}(w_1, w_0)^+$. The marginal effect of a wage increase in period 0, which becomes the worker's reference wage in period 1, on the expected continuation value of the employment relationship to the firm is given by (where $dF \equiv dF(q_1|q_0)$):

$$\begin{aligned}\frac{\partial}{\partial r_1} \int_{\underline{q}(w_0, \lambda)}^{\infty} J_1(w_0, q_1) dF &= \int_{\underline{q}(w_0, \lambda)}^{q^l(w_0, \lambda)} J_{1, r_1}(w_0, q_1)^- dF - \underline{q}_r J_1(w_0, \underline{q}) f(\underline{q}|q_0) \\ &\quad + q_r^l \lim_{\epsilon \rightarrow 0} J_1(w_0, q^l - \epsilon)^- f(q^l|q_0) + \int_{q^l(w_0, \lambda)}^{q^u(w_0)} J_{1, r_1}(w_0, q_1)^= dF \\ &\quad - q_r^l \lim_{\epsilon \rightarrow 0} J_1(w_0, q^l)^= f(q^l|q_0) + q^{u'} \lim_{\epsilon \rightarrow 0} J_1(w_0, q^u)^= f(q^u|q_0) \\ &\quad + \int_{q^u(w_0)}^{\infty} J_{1, r_1}(w_0, q_1)^+ dF - q^{u'} \lim_{\epsilon \rightarrow 0} J_1(w_0, q^u + \epsilon)^- f(q^u|q_0).\end{aligned}$$

By definition, $J_1(w_0, \underline{q}) = 0$, and the continuity of the optimal effort function and wage imply $\lim_{\epsilon \rightarrow 0} J_1(w_0, q^l - \epsilon)^- = J_1(w_0, q^l)^=$ and $\lim_{\epsilon \rightarrow 0} J_1(w_0, q^u + \epsilon)^+ = J_1(w_0, q^u)^=$. Hence, the derivatives with respect to the integral limits q^l and q^u cancel each other out, which

yields:

$$\begin{aligned} \int_{\underline{q}(w_0, \lambda)}^{\infty} J_{1,r_1}(w_0, q_1) dF &= \int_{\underline{q}(w_0, \lambda)}^{q^l(w_0, \lambda)} J_{1,r_1}(w_0, q_1)^- dF + \int_{q^l(w_0, \lambda)}^{q^u(w_0)} J_{1,r_1}(w_0, q_1)^= dF \\ &\quad + \int_{q^u(w_0)}^{\infty} J_{1,r_1}(w_0, q_1)^+ dF \end{aligned}$$

Now, for $q_1 \in [\underline{q}(w_0, \lambda), \infty) \setminus [q^l(w_0, \lambda), q^u(w_0)]$ (i.e. where $w \neq r$):

$$\begin{aligned} J_{1,r_1}(r_1, q_1)^{\pm} &= y_e \tilde{e}_w^{\pm} \tilde{w}_r^{\pm} + y_e \tilde{e}_r^{\pm} - s' \tilde{w}_r^{\pm} \\ &= \tilde{w}_r^{\pm} [y_e \tilde{e}_w^{\pm} - s'] + y_e \tilde{e}_r^{\pm} \\ &= y_e \tilde{e}_r^{\pm} < 0, \end{aligned}$$

since from the first-order condition $y_e \tilde{e}_w^{\pm} - s' = 0$ and we deduced in Theorem 1 that $\tilde{e}_r < 0$.

For $q_1 \in [q^l(w_0, \lambda), q^u(w_0)]$ (i.e. where $w = r$ and optimal effort is \tilde{e}^n):

$$J_{1,r_1}(r_1, q_1)^= = \pi_w = -s'(w_0) < 0.$$

As such,

$$\int_{\underline{q}(w_0, \lambda)}^{\infty} J_{1,r_1}(w_0, q_1) dF = \int_{\underline{q}}^{q^l} y_e \tilde{e}_r^- dF - \int_{q^l}^{q^u} s' dF + \int_{q^u}^{\infty} y_e \tilde{e}_r^+ dF < 0; \quad (14)$$

which corresponds to our definition of $\Phi(w_0, \lambda)$. □

Proof of Theorem 3. The proof is qualitatively similar to the proof of Theorem 2, so the details are largely omitted. Let us, however, dwell on the condition that

$$\Psi_w + \delta \Phi_{w_0} < 0$$

establishing concavity of the value function in w_0 . We know from the proof of Theorem 2 that $\Psi(w; q_0, r_0, \lambda)$ is decreasing in w , as $\Psi_w < 0$ for $w \neq r$ and at $w = r$ there is a jump down. Recalling the expression for $\Phi(w_0, \lambda)$ in (14) and recognising that both the integrand (except in the case of $q_1 \in [q^l(w_0, \lambda), q^u(w_0)]$) and the limits of integration depend on w_0 ,

we deduce that

$$\begin{aligned}\Phi_{w_0} = & \int_{\underline{q}}^{q^l} \frac{d}{dr} \{y_e \tilde{e}_r\} dF - \underline{q}_r y_e \tilde{e}_r f(\underline{q}|q_0) + q_r^l y_e \lim_{\epsilon \rightarrow 0} \tilde{e}_r(w_0 - \epsilon, w_0, \lambda) f(q^l|q_0) \\ & + q_r^l s' f(q^l|q_0) - q^{u'} s' f(q^u|q_0) \\ & - q^{u'} y_e \lim_{\epsilon \rightarrow 0} \tilde{e}_r(w_0 + \epsilon, w_0, \lambda) f(q^u|q_0) + \int_{q^u}^{\infty} \frac{d}{dr} \{y_e \tilde{e}_r\} dF.\end{aligned}$$

Now, from the expressions for \tilde{e}_w and \tilde{e}_r in the proof of Theorem 1 it follows that $\lim_{\epsilon \rightarrow 0} \tilde{e}_r(w_0 \pm \epsilon, w_0, \lambda) = -\lim_{\epsilon \rightarrow 0} \tilde{e}_w(w_0 \pm \epsilon, w_0, \lambda)$. Moreover, when $w \neq r$ the first-order condition holds with equality, which implies that $y_e \tilde{e}_w - s' = 0$. These statements together give us $y_e \lim_{\epsilon \rightarrow 0} \tilde{e}_r(w_0 \pm \epsilon, w_0, \lambda) + s' = 0$, which allows several terms to cancel in the above expression. Noting that $\frac{d}{dr} \{y_e \tilde{e}_r\} = y_e \tilde{e}_{rr} + y_{ee} [\tilde{e}_r]^2$ then allows us to conclude that

$$\Phi_{w_0} = \int_{\underline{q}}^{q^l} [y_e \tilde{e}_{rr} + y_{ee} [\tilde{e}_r]^2] dF + \int_{q^u}^{\infty} [y_e \tilde{e}_{rr} + y_{ee} [\tilde{e}_r]^2] dF - \underline{q}_r y_e \tilde{e}_r f(\underline{q}|q_0).$$

Hence, after noticing that $\tilde{e}_{rr} = -\tilde{e}_{ww}$ and collecting this term as the common factor, we can rewrite the expression $\Psi_w + \delta \Phi_{w_0}$ as:

$$\begin{aligned}\tilde{e}_{ww} \left\{ y_e - \delta \left[\int_{\underline{q}}^{q^l} y_e dF + \int_{q^u}^{\infty} y_e dF \right] \right\} \\ + y_{ee} [\tilde{e}_w]^2 + \delta \left[\int_{\underline{q}}^{q^l} y_{ee} [\tilde{e}_r]^2 dF + \int_{q^u}^{\infty} y_{ee} [\tilde{e}_r]^2 dF \right] \\ - \delta \underline{q}_r y_e \tilde{e}_r f(\underline{q}|q_0). \quad (15)\end{aligned}$$

We know from Theorem 1 that $\tilde{e}_r < 0$ and $\tilde{e}_{ww} < 0$, and from Theorem 2 that $\underline{q}_r > 0$. This implies that: the first line in the expression above is negative only if the term in curly brackets is positive; the second line is negative; and the last line is positive.

The term in curly brackets captures the difference between the ‘current’ and the ‘expected discounted future’ marginal effect of effort on output (i.e. the effect of a greater reciprocity today, due to a higher wage, *versus* lower reciprocity in the future, due to a higher reference wage); while the last line captures the marginal increase in the firm’s lay-off reservation productivity, which reduces the support of the distribution over which

the firm will employ a worker with a higher reference wage in the subsequent employment period. As such we can deduce that if the current effect of reciprocity dominates the expected discounted future effect of reciprocity on the firm's value of the employment relationship, and if the firm's lay-off reservation productivity does not increase too much, then expression (15) will be negative, as required.

Nevertheless, under Assumption D3 this deduction always holds and the proof of the nature of the optimal wage follows the same steps as the proof of Theorem 2 where Ψ is replaced with $\Psi + \delta\Phi$. \square

Proof of Proposition 2. The proof relies on investigation of the first-order condition of the two optimisation problems, noting from Proposition 1 that $\Phi(w_0, \lambda) < 0$. First we show that $\hat{q}^l(r, \lambda, \delta) > \tilde{q}^l(r, \lambda)$. Suppose, by contradiction, that $\hat{q}^l \leq \tilde{q}^l$, then the fact that $\Psi_q > 0$ (see the preliminaries in the proof of Theorem 2) implies

$$0 \equiv \lim_{\epsilon \rightarrow 0} \Psi(r - \epsilon; \tilde{q}^l, r, \lambda) \geq \lim_{\epsilon \rightarrow 0} \Psi(r - \epsilon, \hat{q}^l, r, \lambda),$$

but then since $\Phi(w, \lambda) < 0$ we have that

$$\lim_{\epsilon \rightarrow 0} \Psi(r - \epsilon, \hat{q}^l, r, \lambda) > \lim_{\epsilon \rightarrow 0} \Psi(r - \epsilon, \tilde{q}^l, r, \lambda) + \delta\Phi(w, \lambda) \equiv 0,$$

yielding a contradiction. That $\hat{q}^u(r, \lambda, \delta) > \tilde{q}^u(r)$ is similarly proved.

We now want to compare $\hat{w}(r, q, \lambda, \delta)^{-+}$ with $\tilde{w}(r, q, \lambda)^{-+}$ where both functions are defined. We demonstrate that $\hat{w}(r, q, \lambda, \delta)^- < \tilde{w}(r, q, \lambda)^-$ for all $q < \hat{q}^l(r, \lambda)$. Suppose, by contradiction, that $\hat{w}^- \geq \tilde{w}^-$. Then the fact that $\Psi_w < 0$ (see the preliminaries in the proof of Theorem 2) implies

$$0 \equiv \Psi(\tilde{w}^-; q, r, \lambda) \geq \Psi(\hat{w}^-; q, r, \lambda),$$

but then $\Phi(w_0, \lambda) < 0$ implies

$$\Psi(\hat{w}^-; q, r, \lambda) > \Psi(\tilde{w}^-; q, r, \lambda) + \delta\Phi(r, \lambda) \equiv 0,$$

yielding a contradiction. The proof that $\hat{w}(r, q, \lambda, \delta)^+ < \tilde{w}(r, q, \lambda)^+$ for all $q > \hat{q}^u(r, \lambda, \delta)$ is

similar and so omitted. \square

Proof of Proposition 3. Consider how the optimal wage changes with the degree of loss aversion. Implicit differentiation of the wage setting rule gives

$$\hat{w}_\lambda = -\frac{\Psi_\lambda + \delta\Phi_\lambda}{\Psi_w + \delta\Phi_w}.$$

By Assumption D3 the denominator is negative, and we know from the preliminaries in the proof of Theorem 2 that $\Psi_\lambda = 0$ if $w \geq r$ and $\Psi_\lambda > 0$ if $w < r$. Recalling the definition of $\Phi(w_0, \lambda)$ in (14) and noting that π_w and q^u are independent of λ , we deduce that

$$\begin{aligned} \Phi_\lambda = & \int_{\underline{q}}^{q^l} \frac{d}{d\lambda} \{y_e \tilde{e}_r\} dF - \underline{q}_\lambda y_e \tilde{e}_r f(\underline{q}|q_0) \\ & + q_\lambda^l y_e \lim_{\epsilon \rightarrow 0} \tilde{e}_r(w_0 - \epsilon, w_0, \lambda) f(q^l|q_0) + q_\lambda^l s' f(q^l|q_0) + \int_{q^u}^{\infty} \frac{d}{d\lambda} \{y_e \tilde{e}_r\} dF. \end{aligned}$$

Now, the derivatives with respect to the integral limits cancel out since as we deduced previously $y_e \lim_{\epsilon \rightarrow 0} \tilde{e}_r(w_0 - \epsilon, w_0, \lambda) + s' = 0$. Moreover, $\frac{d}{d\lambda} \{y_e \tilde{e}_r\} = y_{ee} \tilde{e}_\lambda \tilde{e}_r + y_e \tilde{e}_{r\lambda}$ which, according to our deductions in Theorem 1, is equal to zero for wages exceeding the reference wage. As such,

$$\Phi_\lambda = \int_{\underline{q}}^{q^l} [y_{ee} \tilde{e}_\lambda \tilde{e}_r + y_e \tilde{e}_{r\lambda}] dF - \underline{q}_\lambda y_e \tilde{e}_r f(\underline{q}|q_0).$$

In Theorem 1 we concluded that $\tilde{e}_\lambda, \tilde{e}_r, \tilde{e}_{r\lambda} < 0$ and we know from Theorem 2 that $\underline{q}_\lambda > 0$. As such, the sign of Φ_λ remains undetermined, so we cannot sign $\Psi_\lambda + \delta\Phi_\lambda$, but rather conclude that $w_\lambda \gtrless 0 \Leftrightarrow \Psi_\lambda + \delta\Phi_\lambda \gtrless 0$. Note that if the layoff reservation productivity does not increase too much, i.e. $\underline{q}_\lambda y_e \tilde{e}_r f(\underline{q}|q_0)$ is sufficiently small, then $\Phi_\lambda < 0$. \square

Proof of Proposition 4. The reservation productivity governing hiring behaviour in the initial contract is characterised by

$$\hat{q}(r_0, \lambda, \delta) = \max\{0, q_0 : \pi(\hat{w}(r_0, q_0, \lambda, \delta), r_0, q_0, \lambda) + \delta \mathbb{E}[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)] = 0\}.$$

Implicit differentiation reveals

$$\begin{aligned}\frac{d\hat{q}}{d\lambda} &= -\frac{y_e[\tilde{e}_w\hat{w}_\lambda + \tilde{e}_\lambda] - s'\hat{w}_\lambda + \delta\frac{d\mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)]}{d\lambda}}{y_q + y_e\tilde{e}_w\hat{w}_{q_0} - s'\hat{w}_{q_0} + \delta\frac{d\mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)]}{dq_0}} \\ &= -\frac{y_e\tilde{e}_\lambda + \hat{w}_\lambda[y_e\tilde{e}_w - s'] + \delta\frac{d\mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)]}{d\lambda}}{y_q + \hat{w}_{q_0}[y_e\tilde{e}_w - s'] + \delta\frac{d\mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)]}{dq_0}}.\end{aligned}\quad (16)$$

Letting $\pi^{-;=+}$ be the profit function when $w_1 < w_0$; $w_1 = w_0$; $w_1 > w_0$, in which effort is given by $\tilde{e}(w_1, w_0, \lambda)^-; \tilde{e}^n; \tilde{e}(w_1, w_0)^+$, we have

$$\begin{aligned}\frac{d\mathbb{E}_0[J_1]}{d\lambda} &= \int_{\underline{q}}^{q^l} \frac{d\pi^-}{d\lambda} dF - \frac{dq}{d\lambda} \pi^-|_{q=\underline{q}} f(\underline{q}|q_0) + \frac{dq^l}{d\lambda} \pi^-|_{q=q^l} f(q^l|q_0) \\ &\quad + \int_{q^l}^{q^u} \frac{d\pi^=}{d\lambda} dF + \frac{dq^u}{d\lambda} \pi^=|_{q=q^u} f(q^u|q_0) - \frac{dq^l}{d\lambda} \pi^=|_{q=q^l} f(q^l|q_0) \\ &\quad + \int_{q^u}^{\infty} \frac{d\pi^+}{d\lambda} dF - \frac{dq^u}{d\lambda} \pi^+|_{q=q^u} f(q^u|q_0).\end{aligned}$$

Noting that $\pi^-|_{q=\underline{q}} \equiv 0$, and that since $\pi^{-,+}|_{q=q^l} = \pi^=$ and $\pi^{-,+}|_{q=q^u} = \pi^=$, the other effects on the limits of integration cancel out, this reduces to

$$\frac{d\mathbb{E}_0[J_1]}{d\lambda} = \int_{\underline{q}}^{q^l} \frac{d\pi^-}{d\lambda} dF + \int_{q^l}^{q^u} \frac{d\pi^=}{d\lambda} dF + \int_{q^u}^{\infty} \frac{d\pi^+}{d\lambda} dF.$$

Now,

$$\begin{aligned}\frac{d\pi^{-;+}}{d\lambda} &= y_e\tilde{e}_\lambda + y_e\tilde{e}_w\tilde{w}_\lambda + y_e\tilde{e}_r\hat{w}_\lambda - s'\tilde{w}_\lambda \\ &= y_e\tilde{e}_\lambda + y_e\tilde{e}_r\hat{w}_\lambda + \tilde{w}_\lambda[y_e\tilde{e}_w - s'] \\ &= y_e[\tilde{e}_\lambda + \tilde{e}_r\hat{w}_\lambda]\end{aligned}$$

since $y_e\tilde{e}_w - s' = 0$ by the first-order condition. However, within the range of rigidity we have

$$\frac{d\pi^=}{d\lambda} = -s'\hat{w}_\lambda.$$

Note that \hat{w}_λ doesn't depend on q_1 and $\tilde{e}_\lambda = 0$ when the wage exceeds the reference wage.

Then, recalling the expression for $\Phi(w_0, \lambda)$ in (14), we have that

$$\begin{aligned}\frac{d\mathbb{E}_0[J_1]}{d\lambda} &= \int_{\underline{q}}^{q^l} y_e[\tilde{e}_\lambda + \tilde{e}_r \hat{w}_\lambda] dF - \int_{q^l}^{q^u} s' \hat{w}_\lambda dF + \int_{q^u}^{\infty} y_e \tilde{e}_r \hat{w}_\lambda dF \\ &= \int_{\underline{q}}^{q^l} y_e \tilde{e}_\lambda dF + \hat{w}_\lambda \int_{\underline{q}}^{q^l} y_e \tilde{e}_r dF - \hat{w}_\lambda \int_{q^l}^{q^u} s' dF + \hat{w}_\lambda \int_{q^u}^{\infty} y_e \tilde{e}_r dF \\ &= \hat{w}_\lambda \Phi + \int_{\underline{q}}^{q^l} y_e \tilde{e}_\lambda dF.\end{aligned}$$

This allows us to write the expression for the numerator in (16) as

$$\begin{aligned}y_e \tilde{e}_\lambda + \hat{w}_\lambda [y_e \tilde{e}_w - s'] + \delta \frac{d\mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)]}{d\lambda} &= \hat{w}_\lambda [y_e \tilde{e}_w - s' + \delta \Phi] + y_e \tilde{e}_\lambda + \delta \int_{\underline{q}}^{q^l} y_e \tilde{e}_\lambda dF \\ &= y_e \tilde{e}_\lambda + \delta \int_{\underline{q}}^{q^l} y_e \tilde{e}_\lambda dF\end{aligned}$$

since the first-order condition for the initial wage implies $y_e \tilde{e}_w - s' + \delta \Phi = 0$.

Similar deductions allow us to conclude that $\frac{d\mathbb{E}_0[J_1]}{dq_0} = \hat{w}_{q_0} \Phi$, and therefore to write the expression for the denominator in (16) as

$$\begin{aligned}y_q + \hat{w}_{q_0} [y_e \tilde{e}_w - s'] + \delta \frac{d\mathbb{E}_0[J_1(\hat{w}(r_0, q_0, \lambda, \delta), q_1)]}{dq_0} &= \hat{w}_{q_0} [y_e \tilde{e}_w - s' + \delta \Phi] + y_q \\ &= y_q.\end{aligned}$$

As such,

$$\frac{d\hat{q}}{d\lambda} = -\frac{y_e \tilde{e}_\lambda + \delta \int_{\underline{q}}^{q^l} y_e \tilde{e}_\lambda dF}{y_q} > 0$$

since we know from Theorem 1 that when $w < r$, then $\tilde{e}_\lambda < 0$. □

References

- Abdellaoui, M., H. Bleichrodt, and C. Paraschiv (2007). Loss Aversion Under Prospect Theory: A Parameter-Free Measurement. *Management Science* 53(10), 1659–1674.
- Agell, J. and H. Bennmarker (2007). Wage Incentives and Wage Rigidity: A Representative

- View From Within. *Labour Economics* 14(3), 347–369.
- Ahrens, S., I. Pirschel, and D. J. Snower (2015). Path-Dependent Wage Responsiveness. *Kiel Working Papers No. 1977*.
- Akerlof, G. A. (1982). Labor Contracts as Partial Gift Exchange. *Quarterly Journal of Economics* 97(4), 543–569.
- Akerlof, G. A. and J. L. Yellen (1990). The Fair Wage-Effort Hypothesis and Unemployment. *Quarterly Journal of Economics* CV(2), 255–283.
- Altmann, S., A. Falk, A. Grunewald, and D. Huffman (2014). Contractual Incompleteness, Unemployment, and Labour Market Segmentation. *Review of Economic Studies* 81(October 2013), 30–56.
- Babecký, J., P. Du Caju, T. Kosma, M. Lawless, J. Messina, and T. Rõõm (2010, dec). Downward Nominal and Real Wage Rigidity: Survey Evidence from European Firms. *Scandinavian Journal of Economics* 112(4), 884–910.
- Bartling, B. and K. M. Schmidt (2015). Reference Points, Social Norms, and Fairness in Contract Renegotiations. *Journal of the European Economic Association* 13(1), 98–129.
- Baucells, M. and R. K. Sarin (2010). Predicting Utility Under Satiation and Habit Formation. *Management Science* 56(2), 286–301.
- Benigno, P. and L. A. Ricci (2011). The Inflation-Output Trade-Off with Downward Wage Rigidities. *American Economic Review* 101(June), 1436–1466.
- Benjamin, D. J. (2015). A Theory of Fairness in Labour Markets. *Japanese Economic Review* 66(2), 182–225.
- Bewley, T. F. (1999). *Why Wages Don't Fall During a Recession*. London: Harvard University Press.
- Bewley, T. F. (2007). Fairness, Reciprocity, and Wage Rigidity. In P. Diamond and H. Vartiainen (Eds.), *Behavioural Economics and Its Applications*. Princeton University Press.

- Bhaskar, V. (1990). Wage Relativities and the Natural Range of Unemployment. *Economic Journal* 100(400), 60–66.
- Bolton, G. E. and A. Ockenfels (2000). A Theory of Equity, Reciprocity, and Competition. *American Economic Review* 90(1), 166–193.
- Campbell, C. M. and K. S. Kamlani (1997). The Reasons for Wage Rigidity: Evidence From Survey of Firms. *Quarterly Journal of Economics* 112, 759–789.
- Chemin, M. and A. Kurmann (2014). Do Workers Feel Entitled to High Wages? Evidence from a Long-Term Field Experiment. *Working Paper*, 1–44.
- Clark, A. E., D. Masclet, and M. C. Villeval (2010). Effort and Comparison Income: Experimental and Survey Evidence. *ILR Review* 63(3), 407–426.
- Cohn, A., E. Fehr, and L. Goette (2014). Fair Wages and Effort Provision: Combining Evidence from a Choice Experiment and a Field Experiment. *Management Science*.
- Cox, J. C., D. Friedman, and S. Gjerstad (2007). A tractable model of reciprocity and fairness. *Games and Economic Behavior* 59, 17–45.
- Danthine, J.-P. and A. Kurmann (2006). Efficiency Wages Revisited: The Internal Reference Perspective. *Economics Letters* 90(2), 278 – 284.
- Danthine, J.-P. and A. Kurmann (2007, dec). The Macroeconomic Consequences of Reciprocity in Labor Relations. *Scandinavian Journal of Economics* 109(4), 857–881.
- Danthine, J.-P. and A. Kurmann (2010, oct). The Business Cycle Implications of Reciprocity in Labor Relations. *Journal of Monetary Economics* 57(7), 837–850.
- Dickens, W. T., L. Goette, L. Erica, S. Holden, J. Messina, E. Mark, J. Turunen, and M. E. Ward (2007). How Wages Change : Micro Evidence from the International Wage Flexibility Project. *Journal of Economic Perspectives* 21(2), 195–214.
- Driscoll, J. C. and S. Holden (2004, apr). Fairness and Inflation Persistence. *Journal of the European Economic Association* 2(2-3), 240–251.

- Dufwenberg, M. and G. Kirchsteiger (2000). Reciprocity and wage undercutting. *European Economic Review* 44(4), 1069 – 1078.
- Dufwenberg, M. and G. Kirchsteiger (2004, may). A Theory of Sequential Reciprocity. *Games and Economic Behavior* 47(2), 268–298.
- Eliasz, K. and R. Spiegler (2014). Reference Dependence and Labor-Market Fluctuations. *NBER Macroeconomics Annual* 28, 159–200.
- Elsby, M. W. L. (2009). Evaluating the Economic Significance of Downward Nominal Wage Rigidity. *Journal of Monetary Economics* 56(2), 154–169.
- Elsby, M. W. L., R. Michaels, and D. Ratner (2015). The Beveridge Curve: A Survey. *Journal of Economic Literature* 53(3), 571–630.
- Elsby, M. W. L., D. Shin, and G. Solon (2016). Wage Adjustments in the Great Recession and Other Downturns: Evidence from the United States and Great Britain. *Journal of Labor Economics* 34, 249–291.
- Falk, A. and U. Fischbacher (2006, feb). A Theory of Reciprocity. *Games and Economic Behavior* 54(2), 293–315.
- Fehr, E. and L. Goette (2005, may). Robustness and real consequences of nominal wage rigidity. *Journal of Monetary Economics* 52(4), 779–804.
- Fehr, E., L. Goette, and C. Zehnder (2009, sep). A Behavioral Account of the Labor Market: The Role of Fairness Concerns. *Annual Review of Economics* 1(1), 355–384.
- Fehr, E., O. Hart, and C. Zehnder (2011). Contracts as Reference Points—Experimental Evidence. *American Economic Review* 101(April), 493–525.
- Fehr, E., O. Hart, and C. Zehnder (2014). How Do Informal Agreements and Revision Shape Contractual Reference Points? *Journal of the European Economic Association: forthcoming*.
- Fehr, E. and K. M. Schmidt (1999). A Theory of Fairness, Competition, and Cooperation. *Quarterly Journal of Economics* 114(3), 817–868.

- Fongoni, M. (2018a). A Theoretical Note on: Asymmetries in Intensity and Persistence of Reciprocity in Labour Markets. *University of Strathclyde Working Papers No. 1815*.
- Fongoni, M. (2018b). Workers' Reciprocity and the (Ir)Relevance of Wage Cyclicalities for the Volatility of Job Creation. *University of Strathclyde Working Papers No. 1809*.
- Gächter, S. and C. Thöni (2010). Social Comparison and Performance: Experimental Evidence on the Fair Wage-effort Hypothesis. *Journal of Economic Behavior & Organization* 76(3), 531 – 543.
- Gneezy, U. and J. A. List (2006). Putting Behavioral Economics to Work: Testing For Gift Exchange in Labor Markets Using Field Experiments. *Econometrica* 74(5), 1365–1384.
- Hart, O. and J. Moore (2008). Contracts as Reference Points. *Quarterly Journal of Economics* CXXIII(February), 1–48.
- Herweg, F. and K. M. Schmidt (2012). Loss Aversion and Ex Post Inefficient Renegotiation. *University of Munich, Working Paper*.
- Herz, H. and D. Taubinsky (2018). What Makes a Price Fair? An Experimental Study of Transaction Experience and Endogenous Fairness Views. *Journal of the European Economic Association* 16(2), 316–352.
- Holden, S. and F. Wulfsberg (2009). How Strong is the Macroeconomic Case for Downward Real Wage Rigidity? *Journal of Monetary Economics* 56, 605–615.
- Holden, S. and F. Wulfsberg (2014). Wage Rigidity, Inflation, and Institutions. *Scandinavian Journal of Economics* 116(2), 539–569.
- Kahneman, D., J. L. Knetsch, and R. H. Thaler (1986). Fairness as a Constraint on Profit Seeking: Entitlements in the Market. *American Economic Review* 76(4), 728–741.
- Kahneman, D. and R. H. Thaler (1991). Economic Analysis and the Psychology of Utility: Applications to Compensation Policy. *The American Economic Review* 81(2), 341–346.
- Kahneman, D. and A. Tversky (1979). Prospect Theory: An Analysis of Decision under Risk. *Econometrica* 47(2), 263–292.

- Kaur, S. (2018). Nominal Wage Rigidity in Village Labor Markets. *American Economic Review (Forthcoming)*.
- Keynes, J. M. (1936). *The General Theory of Employment, Interest and Money*. London: MacMillan and Cambridge University Press.
- Kőszegi, B. and M. Rabin (2006). A Model of Reference-Dependent Preferences. *Quarterly Journal of Economics CXXI*(November), 1133–1165.
- Koch, C. (2017). Do Reference Points Erode Fairness? Experimental Evidence on Wage Rigidity. *Working Paper*, 1–36.
- Koenig, F., A. Manning, and B. Petrongolo (2014). Reservation Wages and the Wage Flexibility Puzzle. *CEP Discussion Paper No 1319*.
- Kube, S., M. A. Maréchal, and C. Puppe (2013). Do Wage Cuts Damage Work Morale? Evidence from a Natural Field Experiment. *Journal of the European Economic Association 11*(4), 853–870.
- Kurmann, A. and E. McEntarfer (2017). Downward Wage Rigidity in the United States: New Evidence from Administrative Data. *Working Paper*.
- Macera, R. and V. L. te Velde (2018). On the Power of Surprising versus Anticipated Gifts in the Workplace. *Working Paper*.
- Malmendier, U., V. L. Velde, and R. Weber (2014). Rethinking Reciprocity. *Annual Review of Economics 6*, 849–974.
- Mas, A. (2006). Pay, Reference Points, and Police Performance. *Quarterly Journal of Economics CXXI*(3), 783–821.
- McDonald, I. M. and H. Sibly (2001). How Monetary Policy can have permanent real effects with only temporary Nominal Rigidity. *Scottish Journal of Political Economy 48*(5), 532–546.
- Nickell, S. and G. Quintini (2003). Nominal Wage Rigidity and the Rate of Inflation. *The Economic Journal 113*(490), 762–781.

- Pissarides, C. A. (2009). The Unemployment Volatility Puzzle: Is Wage Stickiness the Answer? *Econometrica* 77(5), 1339–1369.
- Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *American Economic Review* 83(5), 1281–1301.
- Rogerson, R. and R. Shimer (2011). Chapter 7 - Search in Macroeconomic Models of the Labor Market . Volume 4, Part A of *Handbook of Labor Economics*, pp. 619 – 700. Elsevier.
- Sliwka, D. and P. Werner (2017). Wage Increases and the Dynamics of Reciprocity. *Journal of Labor Economics* 35(2), 299–344.
- Snell, A., H. Stüber, and J. Thomas (2018). Downward Real Wage Rigidity and Equal Treatment Wage Contracts: Theory and Evidence. *Review of Economic Dynamics* 30, 265–284.
- Snell, A. and J. P. Thomas (2010). Labor Contracts, Equal Treatment, and Wage-Unemployment Dynamics. *American Economic Journal: Macroeconomics* 2(3), 98–127.
- Summers, L. H. (1988). Relative Wages, Efficiency Wages, and Keynesian Unemployment. *American Economic Review* 78(2), 383–388.
- Tversky, A. and D. Kahneman (1991). Loss Aversion in Riskless Choice: a Reference Dependent Model. *Quarterly Journal of Economics* 106(4), 1039–1061.
- Williamson, O. E. (1985). *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: The Free Press.